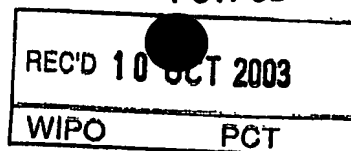




10 / 526479  
PCT/GB 2003 / U U S 8 6 U



INVESTOR IN PEOPLE

GB03 / 03860

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales NP10 8QQ 04 MAR 2005

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

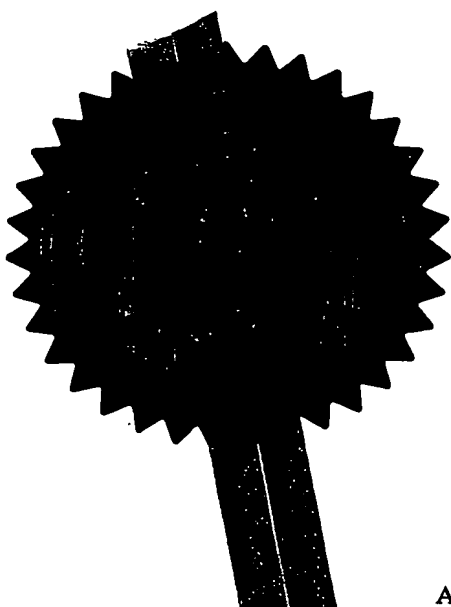
In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 29 September 2003





1/77

# Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)



The Patent Office

Cardiff Road  
Newport  
South Wales  
NP10 8QQ

1. Your reference

P.86234A SER

2. Patent application number

(The Patent Office will fill in this part)

0304521.8

28FEB03 E788455-1 000192  
P01/7700 0.00-0304521.8

3. Full name, address and postcode of the or of each applicant (underline all surnames)

ISOGENICA LIMITED  
Stuart House  
City Road  
Peterborough PE1 1QF

Patents ADP number (if you know it)

8113417001

If the applicant is a corporate body, give the country/state of its incorporation

United Kingdom

4. Title of the invention

DISPLAY LIBRARY

5. Name of your agent (if you have one)

J.A. KEMP & CO.

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

14 South Square  
Gray's Inn  
London  
WC1R 5JJ

Patents ADP number (if you know it)

26001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number  
(if you know it)

Date of filing  
(day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing  
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

Yes

- a) any applicant named in part 3 is not an inventor, or
  - b) there is an inventor who is not named as an applicant, or
  - c) any named applicant is a corporate body.
- See note (d))

# Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document.

Continuation sheets of this form

Description 63

Claim(s) 5

Abstract 1

Drawing(s) 6

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

J.A. KEMP & CO.

Date 27 February 2003

12. Name and daytime telephone number of person to contact in the United Kingdom
- S.E. ROQUES  
020 7405 3292

## Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

## Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

## DISPLAY LIBRARY

### Field of the Invention

The present invention relates generally to recombinant DNA technology and, more particularly, to *in vitro* methods for constructing and screening DNA libraries for DNA sequences that encode biologically active molecules.

### Background of the Invention

Isolating an unknown gene which encodes a desired peptide from a recombinant DNA library can be a difficult task. The use of hybridisation probes may facilitate the process, but their use is generally dependent on knowing at least a portion of the sequence of the gene which encodes the protein. When the sequence is not known, DNA libraries can be expressed in an expression vector, and antibodies have been used to screen plaques or colonies for the desired protein antigen. This procedure has been useful in screening small libraries, but rarely occurring sequences which are represented in less than about 1 in  $10^5$  clones, as is the case with rarely occurring cDNA molecules or synthetic peptides, can be easily missed, making screening libraries larger than  $10^6$  clones at best laborious and difficult. Screening larger libraries has required the development of methods designed to address the isolation of rarely occurring sequences, which are based on the co-selection of molecules, along with the DNAs that encode them. *In vivo* methods have been developed to screen large libraries, such as phage display and "peptides on plasmids" using lacI fusions of peptides.

Phage display is based on DNA libraries fused to the N-terminal end of filamentous bacteriophage coat proteins and their expression in a bacterial host resulting in the display of foreign peptides on the surface of the phage particle with the DNA encoding the fusion protein packaged in the phage particle (Smith G. P., 1985, Science 228: 1315-1317). Libraries of fusion proteins incorporated into phage, can then be selected for binding members against targets of interest (ligands). Bound phage can then be allowed to reinfect *Escherichia coli* (*E. coli*) bacteria and then amplified and the selection repeated, resulting in the enrichment of binding members

(Parmley, S. F., & Smith, G. P. 1988., Gene 73: 305-318; Barrett R. W. *et al.*, 1992, Analytical Biochemistry 204: 357-364 Williamson *et al.*, Proc. Natl. Acad. Sci. USA, 90: 4141-4145; Marks *et al.*, 1991, J. Mol. Biol. 222: 581-597).

LacI fusion plasmid display is based on the DNA binding ability of the lac repressor. Libraries of random peptides are fused to the C-terminal end of the lacI repressor protein. Linkage of the LacI-peptide fusion to its encoding DNA occurs via the lacO sequences on the plasmid, forming a stable peptide-LacI-peptide complex. These complexes are released from their host bacteria by cell lysis, and peptides of interest isolated by affinity purification on an immobilised receptor target. The plasmids thus isolated can then be reintroduced into *E. coli* by electroporation to amplify the selected population for additional rounds of screening (Cull, M. G. *et al.* 1992. Proc. Natl. Acad. Sci. U.S.A. 89:1865-1869).

These bacterial methods are limited by the size of the library that can be created by current methods of introducing DNA into host bacteria, the potential cellular toxicity of the expressed peptides introduced, and by the inability to introduce post-translational modifications, or to incorporate novel amino acids into the expressed peptide.

An entirely *in vitro* ribosome system has been described based on the linkage of peptides to the RNA encoding them through the ribosome (W091/05058). Ribosome display has also been used for the selection of single-chain Fv antibody fragments (scFv) (Matheakis, L. C. *et al.*, 1994 Proc. Natl. Acad. Sci. USA, 91: 9022-9026; Hanes, J. & Pluckthun, A. 1997 Proc. Natl. Acad. Sci. USA, 94: 4937-4942). This method suffers from the lower stability of the RNA genetic material and the increased degradation likely under certain selection conditions where RNase is likely to be present.

The *in vitro* method described by Griffiths and Tawfik (WO 99/02671 and WO 00/40712) addresses some of these concerns by compartmentalizing DNA prior to expression of peptides, which then modify the DNA within the compartment. Peptides capable of modifications, resulting from enzymatic activity of interest, are then selected in a subsequent step. However, no direct selection of peptide binding

activity is possible of both peptide and DNA without modification of the DNA encoding that peptide, and by releasing the modified DNA from the compartment.

Another prior art method, covalent display technology, or CDT, is described in WO9837186. This method relies on covalent linkage of protein to DNA to retain the linkage of genotype to phenotype, through the *cis* action of the crosslinking protein. This method teaches that two requirements are needed for successful use of this technique. Firstly, proteins are required which interact *in vitro* with the DNA sequence which encodes them (*cis* action), and secondly, said proteins must establish a covalent linkage to their own DNA template. This method suffers from the fact that the DNA is chemically modified which can prevent the recovery and identification of the binding peptide of interest.

There remains a need for more versatile *in vitro* methods of constructing peptide libraries in addition to the methods described above, which can allow direct selection of binding activity, as well as for enzymatic activity, and that allow efficient production of complex peptide structures, while still allowing recovery of intact genetic material encoding the peptide of interest.

### Summary of the invention

The present invention therefore provides a method for producing an *in vitro* peptide expression library comprising a plurality of peptides, wherein each peptide is linked to a DNA construct encoding the peptide, comprising the steps of:

- (a) providing a DNA construct comprising:
  - (i) a DNA target sequence;
  - (ii) DNA encoding a library member peptide; and
  - (iii) DNA encoding a peptide capable of non-covalently binding directly or indirectly to said DNA target sequence of (ii);wherein said DNA construct and encoded protein are selected to have *cis*-activity;
- (b) expressing a plurality of DNA constructs according to (a), wherein said DNA constructs encode a plurality of library member peptides

such that each expressed peptide is non-covalently linked to the DNA from which it was produced.

Also provided is a method for producing an *in vitro* peptide expression library comprising a plurality of peptides, wherein each peptide is linked to the DNA construct encoding the peptide, comprising the steps of:

- (a) providing a DNA construct comprising:
  - (i) DNA encoding a library member peptide; and
  - (ii) DNA encoding a peptide capable of non-covalently binding to a bifunctional agent;wherein said DNA construct and encoded protein are selected to have *cis*-activity;
- (b) binding a bifunctional agent or a DNA tag capable of binding a bifunctional agent to said DNA construct of (a), wherein said bifunctional agent is capable of binding to the peptide encoded by said DNA of (ii); and
- (c) expressing a plurality of DNA constructs according to (b), wherein said DNA constructs encode a plurality of library member peptides such that each expressed peptide is linked via said bifunctional agent to the DNA from which it was produced.

The present invention extends to the peptide libraries produced by such methods and to the DNA constructs used in such methods.

The present invention also provides methods of screening *in vitro* peptide expression libraries of the invention. In one aspect there is provided a method of identifying and/or purifying a peptide exhibiting desired properties from an *in vitro* peptide expression library produced according to the method of any one of the preceding claims, comprising at least the steps of (a) screening said library and (b) selecting and isolating the relevant library member. In a second aspect, there is provided a method of identifying a specific ligand binding peptide, said method comprising at least the steps of (a) screening an *in vitro* peptide expression library produced according to the method of the invention with ligand molecules which are optionally bound to a solid support; (b) selecting and isolating a library member

binding to said target molecule; and (c) isolating the peptide which binds specifically to said target molecule. In a third aspect there is provided a method of identifying and/or purifying a peptide having the ability to bind a specific DNA target sequence comprising at least the steps of (a) providing an in vitro expression library according to the invention wherein said peptide or protein of (iii) is a library member peptide having DNA binding activity and wherein said DNA target sequence of (i) is the target sequence of interest; (b) selecting and isolating a library member in which the encoded protein binds to said target sequence; (c) isolating the peptide which binds to said target sequence.

In addition to isolating and/or identifying specific peptides from the libraries of the invention, the screening methods of the invention may be used to isolate and/or identify the DNA encoding specific peptides from the library.

#### **Brief Description of the Figures**

Figure 1 gives a schematic representation of a method by which a DNA construct of the invention may be linked to the peptide that it encodes.

Figure 2 give a schematic representation of a method of the invention by which a DNA binding protein may be converted to a cis-acting DNA binding protein.

Figure 3 gives a schematic representation of how a target sequence specific DNA binding protein may be isolated from a library of the invention.

Figure 4 gives a schematic representation of how a library protein may be linked to its coding DNA through cis action and the use of a bi-specific binding molecule.

Figure 5 shows the specificity of anti-V5 antibody binding clones. ELISA screening, read at 450nm, of the seven clones (1-7) that show specific binding to anti-V5 antibody. The bars in group of four represent the ELISA signal of the clones screened against from left to right; anti-human kappa region antibody, anti-V5 antibody, BSA, and blank. A negative control that neither express CK nor V5 is also presented (8).

Figure 6 shows culture supernatant ELISA OD 450nm signals for peptides recovered after 5 rounds of selection against *B.globigii* spores in Example 4. A. =



clone1e; B. = clone1f; C. = clone1g; D. = clone8a; E. = clone10c; F. = clone10e; G. = negative control.

Figure 7 shows OD 450nm signals for peptides isolated after 4 rounds of selection against anti-V5 antibody in Example 5. A. = P1C12; B. = P2H1; C. = P1B5; D. = P2B8. Peptide-phage were tested against anti-V5 and anti-ACTH peptide antibodies.

### **Brief Description of the Sequences**

SEQ ID Nos 1 to 11, 19 to 23, and 26 to 35 show the primers used in the Examples.

SEQ ID NO: 12 shows the sequence of the TAC-MYC-CK-REPA-CIS-ORI construct, SEQ ID NO: 13 shows the sequence of the TAC-MYC-V5-REPA-CIS-ORI construct, SEQ ID NO: 24 shows the sequence of the TAC-V5-REPA-CIS-ORI-408 construct and SEQ ID NO: 25 shows the sequence of the TAC-NNB-REPA-CIS-ORI-408 construct.

SEQ ID NO: 14 shows the estrogen receptor target recognition sequence.

SEQ ID Nos 15 and 16 show the DNA and amino acid sequences of the repA gene from the R1 plasmid of the incFII incompatibility group. SEQ ID Nos 17 and 18 show the sequences of the CIS DNA element and ori sequence from the same system.

SEQ ID Nos 36 to 39 show the sequences of peptides isolated after selection in Example 5.

### **Detailed Description of the Invention**

The present invention relates to the construction and screening of a library for a nucleotide sequence which encodes a peptide of interest *in vitro*. The constructs encoding the peptide of interest are designed such that the expressed peptide shows cis activity for the construct. Cis activity is defined as the ability of the peptide to bind to the DNA from which the peptide was produced, i.e. from which it was transcribed and translated. *In vitro* expression of the construct results in binding of the peptide to the DNA encoding that same peptide molecule by non-covalent

interaction. This differs from the teaching of WO 98/37186, which does not allow for the possibility of *in vitro* non-covalent interaction between protein and the DNA it encodes, and indeed specifically excludes such interactions from having any practical use for library screening. The present invention has the advantage over the methods of WO 98/37186 that the encoded protein may be separated from the DNA which encoded it without damaging the DNA.

Non-covalent binding refers to an association that may be disrupted by methods well known to those skilled in the art, such as the addition of an appropriate solvent, or a change in ionic conditions, for example, the addition of low pH glycine or high pH triethylamine. In the present case, a typical example of non-covalent binding would be the non-covalent interaction between a DNA binding protein and a DNA molecule. Conversely, when a covalent linkage is formed between the DNA and the encoded polypeptide, the displayed peptide or protein will not be released from the DNA by ionic conditions and solvents that would disrupt non-covalent DNA binding protein:DNA interactions. For example, the bacterial replication protein repA binds non-covalently to its target DNA sequence oriR and can be released from this target DNA sequence at salt concentrations greater than 0.2M KCl (Giraldo R. & Diaz R. 1992 J. Mol. Biol. 228: 787-802). This salt concentration would not affect a covalent linkage, which would require much harsher conditions to release the covalently bound protein, with the increased risk of damage to the recovered DNA.

The current invention describes *cis* activity and non-covalent binding which allow the encoded peptide or protein to remain associated with the DNA construct with a half life sufficient to allow individual peptides and the associated DNA encoding that peptide with an activity of interest to be separated from the resulting mixture of protein DNA complexes. For example, the association between the encoded protein and its DNA may have a half life of up to 30 minutes, up to 45 minutes, up to one hour, up to 2 hours, up to 6 hours or up to 12 hours. The screening methods of the invention may therefore be carried out immediately after construction of the library, or later, for example up to one, up to two, up to six, up to twelve hours or up to twenty four hours later.

Surprisingly, therefore, the invention described herein demonstrates that such encoded peptides or proteins can be expressed *in vitro* and bound to the DNA encoding that peptide in the presence of other DNA sequences. The invention also demonstrates that covalent linkage between protein and DNA is not required to maintain such *cis* activity, and that a non-covalent interaction between DNA and binding protein is sufficient to allow selection of peptides in an *in vitro* expression and selection system. Additionally, the invention demonstrates that a peptide of interest can be selected from a mixture of non-binding peptides, and the DNA encoding that peptide may be recovered.

According to the present invention, individual DNA library members, each of which encodes a peptide to be expressed in the peptide expression library (library member peptide), are placed in a suitable DNA construct. The DNA construct into which the DNA library member is placed includes all the sequences necessary to allow expression of the library member peptide from the construct and to allow the peptide encoded by the construct to bind to the DNA construct which encoded it. Each peptide in the library will typically comprise a fusion protein comprising the library member peptide fused to a peptide involved in binding of the fusion protein to the relevant DNA construct. Such fusion proteins may comprise further sequences and said library peptide may be joined to said binding peptide via a linker sequence.

A plurality of such constructs, encoding a plurality of different library member peptides form a DNA library of the invention. Expressing such a library of DNA molecules results in the non-covalent binding of individual encoded proteins to the DNA which encoded them and from which they have been transcribed and translated, in the presence of many other DNA molecules that encode other members of the library. The sequence encoding the peptide library member present in a particular encoded protein will therefore be present in the DNA which is bound to that protein. This process therefore links the library member peptide, in a biologically active form (usually having a binding activity) to the specific library member DNA sequence encoding that peptide, allowing selection of peptides of interest, for example due to a particular binding activity, and subsequent isolation and identification of the DNA encoding that library member peptide.

For the purposes of the invention a DNA library is therefore a population of DNA constructs. Each construct comprises a DNA sequence encoding a peptide to be expressed as a library member peptide and each contains appropriate promoter, translation start and stop signals. A DNA library of the invention will contain a plurality of such DNA molecules. A plurality of DNA constructs are provided each encoding a library member peptide to provide a plurality of different library members. Preferably a DNA library will contain at least  $10^4$  discrete DNA molecules. For example, a DNA library may contain more than  $10^6$ , more than  $10^8$ , more than  $10^{10}$ , more than  $10^{12}$  or more than  $10^{14}$  discrete DNA molecules.

A peptide expression library is defined as a population of peptide sequences expressed from a library of DNA molecules. A peptide expression library of the present invention therefore encompasses a library of peptides which are non-covalently bound to the DNA which encoded them. For example, a peptide expression library of the present invention may be a library of at least  $10^4$  discrete proteins each comprising a library peptide sequence, expressed from a library of DNA molecules. A peptide expression library of the invention may be any library formed by the expression of a DNA library of the present invention.

A peptide library member can be defined as an amino acid chain of random composition of at least two amino acids in length, or part or all of a naturally occurring protein such as an enzyme, a binding molecule such as an antibody or a fragment thereof.

A DNA construct according to the present invention may comprise DNA encoding a library member peptide and means for the encoded peptide to bind to the encoding DNA construct. In addition to DNA encoding a library member peptide, a suitable DNA construct of the invention comprises at least a DNA target sequence and DNA encoding a peptide capable of binding directly or indirectly to the DNA target sequence.

According to the present invention, the DNA construct and the encoded protein are selected to have cis-activity. That is, the encoded protein has the ability to bind specifically to the DNA molecule which encoded it. For example, cis-activity may function to allow the encoded DNA binding peptide to bind specifically

(directly or indirectly) to the target sequence of the DNA construct which encoded it rather than to the target sequence of another DNA construct.

In some cases, cis activity may be provided by a cis-acting DNA element. In other cases, a separate cis acting DNA element per se may not be required where the nature of the encoding DNA inherently confers cis activity on the encoded peptide.

A cis-acting DNA element may be provided in the DNA construct together with the DNA encoding a peptide that interacts with that cis element. For example, in the case of the cis element from the repA system discussed below, DNA encoding a fragment of the repA sequence comprising at least 20 amino acids from the C terminal of repA may be provided along with the cis DNA element. It may be possible to confer cis activity upon a DNA binding peptide that is not normally cis-acting by including in the DNA construct a cis-acting DNA element and any further sequences necessary for its action. For example, DNA encoding a peptide that interacts with the cis element used may be included in the DNA construct.

Alternatively, a peptide that interacts with the cis element may be part of the DNA binding peptide. For example, the DNA binding peptide may be repA which comprises the sequence that interacts with the repA cis element. Alternatively, the DNA binding peptide may bind to its encoding DNA in cis without the need for a separate cis element.

A suitable cis-acting DNA element may be any element which allows cis-activity. Such a cis-acting DNA element may act, for example, by interacting with the machinery involved in translation and transcription of the DNA construct to delay the production and release of the encoded peptide.

Any DNA element which allows the encoded peptide to bind specifically to the DNA molecule which encoded it may be used as a cis-acting DNA element according to the present invention. One example of a suitable cis-acting DNA element is that of the repA-cis system described in more detail below. In that system, RNA polymerase is paused by loops in the 5' cis sequence prior to the rho dependent termination of transcription. The action of the cis-acting DNA element therefore allows the encoded binding peptide to bind to the DNA target sequence in the construct from which it was produced.

Preferably, the cis DNA element will be located 3' in the DNA construct to the library member peptide and to the peptide or protein capable of binding to the DNA target sequence. This means that these sequences may be transcribed and translated before the RNA polymerase reaches the cis acting sequence.

According to the present invention, the binding peptide may be linked to the DNA construct directly or indirectly. In the case of direct binding, the binding peptide binds directly and non-covalently to the DNA target sequence. In the case of indirect binding, the link between the binding peptide and DNA construct is provided by a further molecule. Such a molecule, for example a bifunctional agent as described below, will associate with both the peptide and the DNA target sequence.

A suitable DNA construct may comprise further sequences, for example suitable promoter sequences to allow expression of the encoded peptide.

One example of a system in which cis-activity exists is the a cis acting incompatibility group plasmid replication protein, termed repA, system. Aspects of this system may be utilised in the present invention as explained below.

Numerous plasmids include sequences encoding repA and cis DNA elements. The repA sequence and cis DNA element present in a DNA construct of the invention may be derived from the same plasmid strain or may be derived from different plasmid strains.

It is believed that the repA-cis system acts as shown in Figure 1. Briefly, RNA polymerase is paused by loops in the 5'-CIS sequence prior to rho dependent termination of transcription. This allows transient C-terminal repA peptide interaction with CIS, and possibly DNA bending. RepA peptide then binds to ori, which is a defined distance away from the terminal amino acid of the repA coding sequence (Prazkier et al. 2000 J. Bacteriology 182: 3972-3980; Prazkier and Pittard 1999 J. Bacteriol. 181: 2765-2772; Masai and Arai. 1988 Nucleic Acids Res. 16: 6493-6514).

The compatibility of a RepA sequence from a plasmid with a cis sequence from another plasmid can be readily determined by monitoring for the interaction of RepA with the selected cis sequence.

Suitable repA proteins and sequences and cis DNA elements include those of the IncI complex plasmids or the IncF, IncB, IncK, IncZ and IncL/M plasmids, which are distantly related at the DNA level, but which control plasmid replication through the action of the cis acting repA protein (Nikoletti et al. 1986 J. Bacteriol. 170:1311-1318; Praskier J. et al. 1991 J. Bacteriol. 173: 2393-2397). Specific plasmids which may be used to provide these sequences include the R1 plasmid of the IncII incompatibility group and the incB plasmid pMU720 (described by Praskier J. & Pittard J. 1999 Role of CIS in replication of an IncB plasmid. J. Bacteriol. 181: 2765-2772). The DNA and amino acid sequences of repA derived from the R1 plasmid of IncII are given in SEQ ID Nos: 15 and 16. The DNA sequence of the cis DNA element from the R1 plasmid of IncII is given in SEQ ID NO: 17. Typically, the cis element is 150 to 200 nucleotides in length. Shorter or larger sequences may be used, so long as the sequence maintains the ability to interact with RepA and display cis activity. Minor variations, such as substitutions or deletions within the cis sequence are also contemplated such as modifications at 1, 5, 10 up to 20 nucleotides within the wildtype cis sequence.

The cis element is required for cis activity of the repA protein (Praskier and Pittard 1999 J. Bacteriol. 181: 2765-2772). The cis DNA element should therefore also be located 3' in the DNA construct to the DNA encoding the repA sequence. On reaching the cis sequence, the RNA polymerase will be paused, allowing the encoded protein to bind the DNA target sequence.

In one embodiment of the present invention, the DNA binding protein itself comprises RepA or a fragment or variant thereof capable of DNA binding, including at least the 20 C-terminal amino acids of RepA capable of binding to the cis DNA element. In this embodiment, the DNA target sequence comprises an ori sequence, for example the oriR sequence. In alternative aspects of the present invention, the DNA binding protein is provided by an alternative protein with the relevant DNA target sequences recognised by such binding protein incorporated into the sequence. In each of these embodiments, DNA-protein binding is direct in that the peptide encoded by the DNA construct will bind directly to the encoding DNA construct. In alternative aspects of the invention, as described in more detail below, the DNA-

protein binding may be indirect through the use of a peptide tag-DNA tag, bifunctional agent and/or suitable linkers.

In one aspect, the same sequence may therefore provide both the peptide capable of binding the DNA target sequence and the C terminal amino acids of repA. Such a sequence may be or may comprise a complete repA sequence, or a fragment or variant thereof of a repA sequence which retains the ability to bind to the DNA target sequence used. Where the repA acts as a DNA binding protein, both cis and ori sequences (Praszkier and Pittard 1999 J. Bacteriol. 181: 2765-2772) are required for cis activity (cis) and DNA binding (ori). In this aspect, therefore, the DNA target sequence is an ori sequence and the peptide or protein capable of binding said target is a repA protein. The position of ori in the DNA constructs of the invention may be varied. As described earlier, suitable repA, cis and ori sequences may be provided by one or more plasmids. For example, suitable sequences may be provided from the IncI complex plasmids or the IncF, IncB, IncK, IncZ and IncL/M plasmids. The DNA sequence of the ori from the R1 plasmid of IncII is given in SEQ ID NO: 18. This sequence, or a fragment thereof may be included in a DNA construct of the invention. A DNA construct of the invention may include a complete ori sequence or may include a fragment thereof which is capable of being bound by the repA protein being used.

The RepA protein used in accordance with the present invention may also comprise a fragment or variant of RepA, so long as such variant or fragment of RepA maintains the ability to bind to the selected ori sequence. Such variant or fragment of RepA may include substitutions, for example, at 1, 2, 3 up to 20 amino acids within the RepA sequence so long as such variants maintain the ability to bind to the ori sequence. A suitable fragment of RepA is an ori binding sequence of RepA. Ori sequences include those which are present in wild type plasmids as described above. Typically, such an ori sequence is 170 to 220 nucleotides in length. Fragments and variants of wild type ori sequences may also be used, so long as such ori sequences maintain the ability to be recognised by RepA. Suitable ori sequences for use in combination with selected RepA proteins can readily be determined by monitoring for the interaction of RepA with such an ori sequence.



The basic principle of the invention may therefore be described with reference to the repA/cis/ori system, as shown in Figure 1. This shows an example of a DNA construct of the invention. This construct comprises, from 5' to 3', a promoter sequence, a sequence encoding a library member peptide, a sequence encoding a repA protein, a cis DNA element and an ori sequence. Briefly, the DNA sequence is transcribed from the promoter by RNA polymerase to RNA. The rho dependent termination function present in the cis DNA element causes the RNA polymerase to pause at this part of the sequence. This allows the repA protein and the library peptide to be translated. The repA protein is then able to bind to the ori sequence, linking the encoded protein to the encoding DNA construct.

In one preferred embodiment, library member DNA sequence(s) are fused to the repA, cis and ori DNA of the IncFII plasmid R1 (Masai H et al. 1983 Proc Natl Acad Sci USA 80: 6814-6818). In this embodiment, the library member DNA sequence(s) of interest may be joined by a region of DNA encoding a flexible amino acid linker, to the 5'-end of the repA DNA, under the control of an appropriate promoter and translation sequences for *in vitro* transcription/translation. Many suitable promoters are known to those skilled in the art, such as the araB, tac promoter or the T7, T3 or SP6 promoters, amongst others. The promoter should be upstream of the polypeptide sequence to be expressed.

The repA family of proteins is used herein by way of example, not limitation. Other unrelated non-covalently binding cis acting DNA binding proteins could be used in this invention.

In a further embodiment, non-cis acting DNA binding proteins may be converted to having cis-activity (see Figure 2). This may be achieved by using such proteins, capable of binding the DNA target sequence, either directly or indirectly, in combination with sequences which can confer cis-activity upon them. Cis activity may be conferred on a binding protein that does not normally act in cis by including in the DNA construct a cis-acting DNA element such as the cis element of the repA system. Such an element may be included to ensure that the DNA binding by the DNA binding protein is cis, that is, an encoded DNA binding protein will bind to the DNA construct from which it has been transcribed and translated.

In one embodiment, a suitable DNA construct may therefore comprise the cis-acting DNA element from the repA system. Such an element may further comprise DNA encoding a portion of the C-terminal end of RepA, preferably at least 20 amino acids, more preferably 30 amino acids, up to 40, 50, 60 or 70 amino acids from the C-terminal portion of repA., wherein said fragment of repA is capable of interacting with the cis-acting DNA element within the construct. DNA encoding sequences of the present invention may comprise wild type sequences encoding the desired fragment of RepA, degenerate sequences encoding fragments of wild type RepA or sequences encoding variants of such fragments of RepA which maintain the ability to interact with the cis element incorporated into the DNA construct. Such variants may include substitution of 1, 2, 3 or 4 amino acids within the 20 amino acid C-terminal of RepA.

The repA family of proteins is used herein by way of example, not limitation. Any cis-acting DNA element capable of conferring cis-activity on a non-cis acting protein could be used .

Any non-cis acting protein may be converted in this way. By way of example, not exclusion, the estrogen receptor DNA binding domain (DBD) can be converted into a cis acting DNA binding protein. The oestrogen receptor DNA binding domain fragment (amino acids 176-282) has been expressed in *E. coli* and shown to bind to the specific double stranded DNA oestrogen receptor target HRE nucleotide sequence, with a similar affinity (0.5nM) to the parent molecule (Murdoch et al. 1990, Biochemistry 29: 8377-8385; Mader et al., 1993, DNAs Research 21: 1125-1132). In one embodiment, the DNA encoding this sequence is fused, preferably at the 3'-end, to the DNA encoding at least the last 20 amino acids of repA, the cis DNA element, and the DNA up to the ori sequence followed by the estrogen receptor target recognition sequence (5'-TCAGG TCAGA GTGAC CTGAG CTAAC ATAAC ACATT CAG-3', SEQ ID NO: 14) which replaces the repA ori recognition sequence. The DNA sequence(s) of interest may then be joined by a region of DNA encoding a flexible amino acid linker, to the 5'-end of to the estrogen receptor DNA fragment, under the control of an appropriate promoter and translation sequences for *in vitro* transcription/translation. Expression of this

polypeptide directs the estrogen receptor DBD to its target sequence, present in place of the normal ori sequence, on the DNA encoding that polypeptide. Protein-DNA complexes can then be isolated by capture on a target protein. Unbound protein-DNA complexes can be washed away, allowing enrichment for DNA encoding peptides or proteins of interest, which can then be recovered by PCR, and enriched further by performing several further cycles of *in vitro* expression and protein-DNA complex capture using methods described previously.

It will be clear that this approach will apply to other DNA binding proteins simply by using the cis DNA element and a sequence encoding at least the C-terminal 20 amino acids of repA, or equivalent elements from a different cis-acting system in the DNA constructs.

In another embodiment, libraries of randomized DNA binding proteins, such as zinc finger proteins, helix-loop-helix proteins or helix-turn-helix proteins by way of example, may be screened for specific binding to a target sequence of interest (see Figure 3). In this embodiment, the ori recognition sequence of repA may be replaced by a target sequence of interest, and the majority of the repA coding sequence by a library of randomised zinc finger proteins. The DNA binding proteins therefore act as both the library member peptides and the proteins capable of binding the DNA target sequence in this aspect. The DNA encoding each zinc finger protein, may additionally be joined, at the 5'-end, to a peptide tag sequence which can be recognized by an another capture protein such as an antibody, and at the 3'-end, to the DNA encoding at least the last 20 amino acids of repA, the cis DNA element, and the DNA up to the ori sequence followed by the target sequence of interest. Expression of this polypeptide directs the zinc finger protein to the target sequence of interest, present in place of the normal ori sequence, on the DNA encoding that polypeptide. Binding to the target sequence will only occur if the randomised zinc finger domain is capable of binding to the sequence of interest. Protein-DNA complexes can then be isolated by capture with a binding protein which recognizes the peptide tag at the N-terminus of the fusion protein polypeptide. Unbound DNA can be washed away, allowing enrichment for DNA encoding zinc finger proteins capable of binding the target sequence, which can then be recovered by PCR, and

enriched further by performing several further cycles of *in vitro* expression and protein-DNA complex capture.

As explained above, the binding peptide may bind directly to the DNA target sequence, for example in the case of a DNA binding protein-target sequence pair, or it may bind indirectly to the DNA target sequence, for example via a bifunctional agent and optionally a DNA tag (see Figure 4):

In one embodiment, DNA encoding a peptide tag which is not able to bind directly to the DNA target sequence is joined to the 5'-end of library member DNA sequence(s) of interest, optionally by a region of DNA encoding a flexible amino acid linker, under the control of an appropriate promoter and translation sequences for *in vitro* transcription/translation. This forms the DNA encoding the binding peptide, as the encoded peptide is linked indirectly to the DNA target sequence. Optionally at the 3'-end of the library member DNA sequence is the DNA encoding at least the last 20 amino acids of repA and the cis DNA element, but not the ori target sequence of repA. The DNA target sequence may be or may comprise a DNA tag. Such a DNA tag may be a single modified base. For example, when preparing the library DNA construct containing the elements described, the DNA may be tagged at the 3'-end with, by way of example not limitation, molecules such as fluorescein or biotin.

Prior to *in vitro* expression, the library DNA fragments may be mixed with a bifunctional agent, one function of which is to recognize and bind to the target sequence which may be at the 5' end of the DNA, in a ratio of one DNA fragment: one bifunctional molecule. The other functional element of this bifunctional agent is a binding agent that can recognize and bind to the peptide tag which may be encoded at the 5'-end of the DNA fragment. By way of example not exclusion, the bifunctional agent can be composed of an Fab fragment recognizing the fluorescein or biotin tag on the DNA, and another Fab fragment recognizing the peptide tag encoded in the DNA. It is clear to those skilled in the art that this bifunctional agent can be made by many different methods such as chemically cross-linking the two elements, or by expressing the two elements as a fusion protein, or as a bi-specific

antibody. Said methods of creating a bifunctional agent are given by way of example not exclusion.

The bifunctional agent may be bound to the DNA construct prior to expression of the encoded peptide or may be provided during expression.

The fusion protein is then transcribed and translated from the DNA construct while bound to the bifunctional agent. The peptide tag is translated first, and can be bound by the second element of the bifunctional agent, prior to release of messenger RNA or RNA polymerase from the DNA. This creates a functional protein-DNA complex where both expressed polypeptide and DNA encoding that peptide are linked through the bifunctional agent. The peptide tag molecule is therefore linked indirectly, but specifically, to the DNA target (tag). By linking the protein to the DNA construct in this way, it is possible to screen for a protein having particular properties, as described below, and then to identify the encoding DNA which is linked to that protein. By using a bifunctional agent rather than covalent binding between the protein and DNA, the DNA construct may be more easily separated from the protein without the risk of damaging the DNA.

Protein-DNA complexes can then be isolated by capture of a target protein. Unbound protein-DNA complexes can be washed away, allowing enrichment for DNA encoding peptides or proteins of interest, which can then be recovered by PCR, and enriched further by performing several further cycles of *in vitro* expression and protein-DNA complex capture using methods described previously.

Additionally, under this embodiment, the DNA can be bound directly, for example by covalent binding, to a bifunctional agent such as a polymer. Such a polymer can contain more than one binding element that could recognise the peptide tag, allowing multivalent display of a peptide expression library molecule in a unit containing the DNA encoding the displayed peptide. By way of example, not limitation, said polymers can be composed of polyethylene as well as other polymeric compounds, capable of being fused to DNA. The DNA construct of the invention may therefore be provided bound to such a bifunctional agent, or bound to a DNA tag as described above which is capable of being bound by such a bifunctional agent.

In all embodiments of the invention, the DNA constructs include appropriate promoter and translation sequences for *in vitro* transcription/translation. Any suitable promoter can be used, such as the ara B, tac promoter, T7, T3 or SP6 promoters amongst others. The promoter is placed so that it is operably linked to the DNA sequences of the invention such that such sequences are expressed.

The DNA encoding the library member peptides may be produced by any suitable means. In particular, such DNA may comprise DNA isolated from cDNA, obtained by DNA shuffling, and synthetic DNA.

The DNA construct may also encode amino acid linkers within the expressed fusion protein. In particular, a flexible amino acid linker may be included to join the DNA binding peptide/RepA to the library member peptide.

According to the invention, with reference to this preferred embodiment, peptide or protein expression libraries, linked to the DNA encoding them, can be generated and peptides with the desired activity selected by the following steps:

*Constructing a library of fusion proteins.*

A DNA library of peptides or proteins may be fused to DNA encoding a peptide capable of binding to the DNA target sequence, such as a *cis* acting DNA binding protein DNA, by a region of DNA encoding a flexible amino acid linker, under the control of an appropriate promoter and with a translation, or ribosome binding site, start and stop codons, in a manner suitable for *in vitro* expression of the peptide library members and binding proteins. In the example of the repA protein, the DNA (such as DNA) library members are fused to the repA DNA binding protein DNA, or a fragment thereof. The *cis* and *ori* sequences may be included in the construct downstream of the other elements. In the case of a DNA library, said DNA constructs are designed to be suitable for *in vitro* transcription and translation.

*Expression and cis binding of DNA library fusion proteins.*

In order to allow *cis* activity, a coupled bacterial transcription/translation environment such as the S30 extract system (Zubay, G. 1973. *Ann. Rev. Genet.* 7: 267) may be used. Expression of the peptide, such as the DNA library member peptide-repA fusion protein, in this environment, will result in binding of the fusion protein to the DNA encoding that fusion protein, provided that both *cis* and *ori*

sequences are present. When libraries of peptide-repA fusion proteins are expressed in this manner, this process results in the production of libraries of protein-DNA complexes where the protein attached to the DNA is encoded by that fragment of DNA from which it was expressed, thereby allowing subsequent selection of both peptides or protein of interest, and the DNA encoding said peptides. The complexity of these libraries is enhanced by the *in vitro* nature of the method, libraries of at least  $10^{10}$ - $10^{14}$  DNA fragments, if not even larger libraries, can easily be generated.

*Selection of the peptide of interest..*

An *in vitro* peptide expression library produced by a method of the present invention may be used to screen for particular members of the library. For example, the library may be screened for peptides with a particular activity or a particular binding affinity. Protein-DNA complexes of interest may be selected from a library by, for example, affinity or activity enrichment techniques. This can be accomplished by means of a ligand specific for the protein of interest, such as an antigen if the protein of interest is an antibody. The ligand may be presented on a solid surface such as the surface of an ELISA plate well, or in solution, for example, with biotinylated ligand followed by capture on to a streptavidin coated surface or magnetic beads, after a library of protein-DNA complexes had been incubated with the ligand to allow ligand-ligand interaction. Following either solid phase or in solution incubation, unbound complexes are removed by washing, and bound complexes isolated by disrupting ligand-ligand interactions by altering pH in the well, or by other methods known to those skilled in the art such as protease digestion, or by releasing the DNA directly from the complexes by phenol-chloroform extraction to denature the repA-ori DNA binding. Recovering bound complexes, reamplifying the bound DNA, and repeating the selection procedure provides an enrichment of clones encoding the desired sequences, which may then be isolated for sequencing, further cloning and/or expression. For example, the DNA encoding the peptide of interest may be isolated and amplified by, for example PCR. In one embodiment, repeated rounds of selection and DNA recovery may be facilitated by the use of sequential nesting of PCR primers. DNA ends are generally damaged after multiple PCR steps. To recover DNA from such damaged molecules required the

primers to be annealed away from the ends of the DNA construct, thereby sequentially shortening the construct with every round of selection.

In one aspect, the DNA construct and/or the encoded protein may be configured to include a tag. Such a peptide or DNA tag, for example as described above, may be used in the separation and isolation of a library member of interest. Such a tag may also be used to hold the library members, for example on a solid support for use in the screening methods described herein.

It can therefore be seen that the screening methods of the present invention may include the further step of selecting and isolating the relevant library member peptide, allowing the peptide exhibiting the desired properties, and also the DNA encoding that peptide, to be identified and purified.

Numerous types of libraries of peptides fused to the cis acting DNA-binding protein can be screened under this embodiment including:

- (i) Random peptide sequences encoded by synthetic DNA of variable length.
- (ii) Antibodies or antibody fragments, for example single-chain Fv antibody fragments. These consist of the antibody heavy and light chain variable region domains joined by a flexible linker peptide to create a single-chain antigen binding molecule.
- (iii) Random cDNA fragments of naturally occurring proteins isolated from a cell population containing an activity of interest.
- (iv) Random peptide sequences inserted into, or replacing a region of a known protein, whereby the known protein sequence acts as a scaffold, which constrains the random peptide sequence. Many such scaffolds have been described, by way of example, not exclusion, CTLA-4 (WO 00/60070), has been used as a scaffold for peptide libraries.

In another embodiment the invention concerns methods for screening a DNA library whose members require more than one chain for activity, as required by, for example, antibody Fab fragments for ligand binding. In this embodiment heavy or light chain antibody DNA is joined to a nucleotide sequence encoding a DNA binding domain of, for example, repA. Typically the unknown antibody DNA library sequences for either the heavy (VH and CH1) or light chain (VL and CL) genes are



inserted in the 5' region of the repA DNA, behind an appropriate promoter and translation sequences. Thus, repA fused to a DNA library member-encoded protein is produced bound to the DNA encoding that protein. The second known chain, encoding either light or heavy chain protein, is expressed separately either:

(a) from the same DNA fragment containing the repA and the first polypeptide fusion protein library, or

(b) from a separate fragment of DNA present in the *in vitro* transcription/translation reaction.

The known chain associates with the library of unknown fusion proteins that are fused to the repA protein and thereby bound to the DNA for the unknown chain. The functional Fab library can then be selected by means of a ligand specific for the antibody.

In order that the invention is more fully understood, embodiments will now be described in more detail by way of example only and not by way of limitation with reference to the figures below.

Examples of some of the embodiments of the invention are given below:

### **Materials and Methods**

The following procedures used by the present applicant are described in Sambrook, J., et al., 1989 supra.: analysis of restriction enzyme digestion products on agarose gels, DNA purification using phenol/chloroform stock solutions, preparation of phosphate buffered saline.

General purpose reagents were purchased from SIGMA-Aldrich Ltd (Poole, Dorset, U.K.). Oligonucleotides were obtained from SIGMA-Genosys Ltd (Cambridgeshire, U.K.). Amino acids, and S30 extracts were obtained from Promega Ltd (Southampton, Hampshire, U.K.). Deep Vent and Taq DNA polymerases were obtained from New England Biolabs (Cambridgeshire, U.K.). Taqplus DNA polymerase was obtained from Stratagene Inc. (Amsterdam, Netherlands). GeneClean DNA gel purification kits were obtained from BIO101 (La Jolla, California, U.S.A.), anti-human Igk antibodies from Immunologicals Direct Ltd

(Oxfordshire, U.K.), anti-c-myc polyclonal from Vector Labs Inc (Cambridgeshire U.K.), and anti-V5 antibody from Abcam Ltd (Cambridgeshire U.K.). Superblock blocking agent was obtained from Perbio Science (Cheshire, U.K.).

### **Example 1. Isolation of specific cis acting protein-DNA complexes**

The *in vitro* expression constructs were prepared by sequentially adding the TAC promoter, the c-myc epitope, either the human kappa constant region or the V5 epitope to the RepA-CIS-ORI region, by PCR amplification. Such constructs can be prepared by many methods known to one skilled in the art, for example, by amplifying different fragments of DNA followed by assembly PCR. In this example, the initial amplification template was the R1 plasmid which contains the RepA-CIS-ORI region (Masai, H. and Arai, K.(1988). DNAs Res. 16, 6493-6514).

(a) Primary amplification. The RepA-CIS-ORI region was PCR amplified from a single colony of the strain ECO K12 harbouring plasmid R1 using 12.5pmol of each of the primers REPAFOR (SEQ ID 01) and ORIREV (SEQ ID 02) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units Taqplus Precision DNA polymerase, 1x PCR reaction buffer (Stratagene Inc, Amsterdam, Netherlands). The REPAFOR primer anneals to the 5'-end of the RepA coding region. The ORIREV primer anneals to the 3'-end of the non-coding ORI region.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 4 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 45 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(b) Secondary amplification. One µl (500 pg) of 100 times diluted gel-purified primary reaction product was re-amplified using 12.5pmol of each of the primers CKREPFOR (SEQ ID 03) and ORIREV (SEQ ID 02) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units Taqplus Precision DNA polymerase, and 1x PCR reaction buffer (Stratagene Inc, Amsterdam, Netherlands). The CKREPFOR primer anneals to the 5'-end of the primary reaction product and appends the 3' part

of the kappa constant region DNA. The ORIREV primer anneals to the 3'-end of the primary reaction product.

---

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 2 minutes, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(c) Third amplification. One µl (500 pg) of 100 times diluted gel-purified primary reaction product was re-amplified using 12.5pmol of each of the primers V5REPFOR (SEQ ID 04) and ORIREV (SEQ ID 02) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units Taqplus Precision DNA polymerase, and 1x PCR reaction buffer (Stratagene Inc, Amsterdam, Netherlands). The V5REPFOR primer anneals to the 5'-end of the primary reaction product and appends the 3' part of the V5 epitope DNA. The ORIREV primer anneals to the 3'-end of the primary reaction product.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 2 minutes, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(d) Fourth amplification. One µl (500 pg) of 100 times diluted pCKV5 plasmid using 12.5pmol of each of the primers MYCCKFOR (SEQ ID 05) and CKREV (SEQ ID 06) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units Taqplus Precision DNA polymerase, and 1x PCR reaction buffer (Stratagene Inc, Amsterdam, Netherlands). The pCKV5 plasmid contains the human kappa constant region cDNA (McGregor DP, Molloy PE, Cunningham C, & Harris WJ. 1994 Mol. Immunol. 31: 219-26) and the V5 epitope DNA (Southern JA, Young DF, Heaney F, Baumgartner WK, Randall RE. 1991 J. Gen. Virol. 72: 1551-7). The MYCCKFOR primer anneals to the 5'-end of the kappa constant region DNA and appends the 3'

---

part of the MYC epitope DNA. The CKREV primer anneals to the 3'-end of the kappa constant region DNA.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 2 minutes, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(e) Fifth amplification. One µl (500 pg) of 100 times diluted pCKV5 plasmid using 12.5pmol of each of the primers MYCV5FOR (SEQ ID 07) and V5REV (SEQ ID 08) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units Taqplus Precision DNA polymerase, and 1x PCR reaction buffer (Stratagene Inc, Amsterdam, Netherlands). The MYCV5FOR primer anneals to the 5'-end of the V5 epitope DNA and appends the 3' part of the MYC epitope DNA. The V5REV primer anneals to the 3'-end of the V5 epitope DNA.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 30 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(f) Sixth amplification. One µl (500 pg) of 100 times diluted pTACP2A plasmid (ref) using 12.5pmol of each of the primers TAC3 (SEQ ID 09) and MYCTACREV (SEQ ID 10) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units Taqplus Precision DNA polymerase, and 1x PCR reaction buffer (Stratagene Inc, Amsterdam, Netherlands). The TAC3 primer anneals to the 5'-end of the TAC promoter DNA. The MYCTACREV primer anneals to the 3'-end of the TAC promoter DNA and appends the 5' part of the MYC epitope DNA.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45

seconds; 72°C, 30 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(g) First assembly PCR. One µl (50 ng) of each of the reaction products in (f) and (d) using 50 pmol of each of the primers TAC5 (SEQ ID 11) and CKREV (SEQ ID 06) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TAC5 primer anneals to the 5'-end of the reaction product (f) and adds 20 nucleotides. The CKREV primer anneals to the 3'-end of the reaction product (d).

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 45 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(h) Second assembly PCR. One µl (50 ng) of each of the reaction products in (f) and (e) using 50 pmol of each of the primers TAC5 (SEQ ID 11) and V5REV (SEQ ID 08) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TAC5 primer anneals to the 5'-end of the reaction product (f) and adds 20 nucleotides. The V5REV primer anneals to the 3'-end of the reaction product (e).

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 45 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(i) Third assembly PCR. One  $\mu$ l (50 ng) of each of the reaction products in (b) and (g) or using 50 pmol of each of the primers TAC3 (SEQ ID 09) and ORIREV (SEQ ID 02) in a 50 $\mu$ l reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TAC3 primer anneals 20 nucleotides downstream to the 5'-end of the reaction product (g). The ORIREV primer anneals to the 3'-end of the reaction product (b). The reaction product in (i) is called TAC-MYC-CK-REPA-CIS-ORI (SEQ ID 12).

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 1 minute, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40 $\mu$ l sterile water using a GeneClean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(j) Fourth assembly PCR. One  $\mu$ l (50 ng) of each of the reaction products in (b) and (h) or using 50 pmol of each of the primers TAC3 (SEQ ID 09) and ORIREV (SEQ ID 02) in a 50 $\mu$ l reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TAC3 primer anneals 20 nucleotides downstream to the 5'-end of the reaction product (g). The ORIREV primer anneals to the 3'-end of the reaction product (b). The reaction product in (i) is called TAC-MYC-V5-REPA-CIS-ORI (SEQ ID 13).

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 2 minutes and 15 seconds at 94°C followed by 30 cycles of 94°C, 45 seconds; 60°C, 45 seconds; 72°C, 1 minute, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40 $\mu$ l sterile water using a GeneClean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

Preparation of *in vitro* transcription/translation reaction. The reaction was set up on ice, using a Promega bacterial linear template S30 coupled *in vitro* transcription/translation reaction kit as follows:

20µl TAC-MYC-CK-REPA-CIS-ORI template (0.5µg of final construct DNA SEQ ID 012 above); 20µl TAC-MYC-V5-REPA-CIS-ORI template (0.5µg of final construct DNA SEQ ID 013 above); 20µl complete amino acid mix (Promega); 80µl S30 Premix; 60µl S30 mix;

and the reaction was allowed to proceed at 25°C for 30 minutes and placed on ice, then diluted 10 fold with blocking buffer (Superblock (Perbio Ltd), 0.1 % Tween 20, 200µg/ml herring sperm DNA).

DNA-protein complex capture. NUNC star immunotubes were coated with 10µg/ml of either anti-c-myc antibody, anti-V5 antibody, or anti-human kappa chain antibody, in 500µl PBS per tube overnight at 4°C. An additional tube was left blank as a negative control. Tubes were washed 2x PBS and blocked for 1 hour at room temperature with Superblock/PBS/0.1mg/ml herring sperm DNA/ 0.1% Tween 20 and then washed 2x PBS. 500µl of diluted transcription/translation reaction was added to each tube and incubated at room temperature for 1 hour. Tubes were washed 5x PBS/0.1% Tween 20, then 1x 30 minutes with 2ml Superblock/PBS/0.1mg/ml herring sperm DNA/ 0.1% Tween 20, then 5x PBS. DNA was recovered with 300µl T.E. buffer plus 300µl phenol/chloroform for 5 minutes with shaking. This was centrifuged at 13,200g for 5 minutes and DNA precipitated with 0.5 volume of 7.5M ammonium acetate, 20µg glycogen and three volumes of absolute ethanol. Following centrifugation, pellets were washed with 70% ethanol, vacuum dried and resuspended in 20µl water. 10µl of recovered DNA was reamplified in 50µl reactions with TAC3 (SEQ ID 09) and ORIREV (SEQ ID 02) primers. Reaction products were electrophoresed on a 1% agarose/TAE gel (Figure 1).

### **Example 2. Separating the RepA-DNA complex**

The two *in vitro* expression constructs (SEQID12 and SEQID13) already described in example 1 were used in a selection experiment against anti-human C-

kappa antibody as described in Example 1, except that DNA was recovered and released from RepA by using either of following methods; Glycine, Triethylamine, Phenol/Chloroform, Proteinase K, and EDTA. These methods are described below.

Glycine: tube was incubated with 500µl of 200mM Glycine, 150mM NaCl (pH2.0) for 10 minutes. The glycine eluate was then transferred to a fresh eppendorf tube and 50µl of 2M Tris (pH 8.5) added.

Triethylamine: the tube was incubated 500µl of 0.1M Triethylamine for 10 minutes and the triethylamine eluate was then transferred to a fresh eppendorf tube and 250µl of 1M Tris (pH 7.4) added.

Phenol/Chloroform: as example 1 above.

Proteinase K: the tube was incubated with 500µl of 100mM Tris (pH 8.0), 10 mM EDTA (pH 8.0), 0.5% SDS for 30 minutes at 37°C. The Proteinase K eluate was then transferred to a fresh eppendorf tube.

EDTA: the tube was incubated with 250µl of 10mM Tris (pH 8.0), 1 mM EDTA 500mM NaCl and 250µl of Phenol/Chloroform for 5 minutes. The EDTA eluate was then transferred to a fresh eppendorf tube.

After recovery of DNA the DNA was Phenol/Chloroform extracted, where appropriate, followed by Ethanol precipitation as described in Example 1. 10ul of resuspended DNA was reamplified in 50ul reactions with TAC3 (SEQID09) and CISREV (SEQID019) primers. The CISREV primer anneals 196 bases upstream of the binding site of ORIREV (SEQID02). Reaction products were electrophoresed on a 1% agarose/TAE gel (data not shown). Only the CK-DNA containing construct (SEQID 12) was amplified, in approximately equivalent amounts.

This not only tells us that any of the methods described above for recovering and releasing DNA from RepA can be used, but this result also suggests that RepA interacts in a non-covalent manner with its cognate DNA.

### **Example 3. Detection of specific anti-V5 binders in a V5-spiking experiment using CIS display technology.**

The *in vitro* expression constructs were prepared by adding the TAC promoter and either the V5 epitope or a 12-mer NNB library to the RepA-CIS-ORI



region, by PCR amplification. Such constructs can be prepared by many methods known to one skilled in the art, for example, by amplifying different fragments of DNA followed by assembly PCR. In this example, the initial amplification template was the R1 plasmid which contains the RepA-CIS-ORI region (Masai, H. and Arai, K.(1988). *Nucleic Acids Res.* 16, 6493-6514).

(a). Primary amplification. The RepA-CIS-ORI region was PCR amplified from a single colony of the strain ECO K12 harbouring plasmid R1 using 12.5pmol of each of the primers REPAFOR (SEQ ID 01) and ORIREV408 (SEQ ID 20) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The REPAFOR primer anneals to the 5'-end of the RepA coding region. The ORIREV408 primer anneals to the downstream of the 3'-end of the non-coding ORI region.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 4 minutes and 30 seconds of 94°C followed by 25 cycles of 94°C, 30 seconds; 60°C, 45 seconds; 72°C, 1 minute, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(b). Secondary amplification. One µl (500 pg) of 100 times diluted gel-purified primary reaction product was re-amplified using 12.5pmol of each of the primers V5(NNB)REPFOR (SEQ ID 21) and ORIREV408 (SEQ ID 20) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The V5(NNB)REPFOR primer anneals to the 5'-end of the primary reaction product and appends the V5 epitope DNA. The ORIREV408 primer anneals to the 3'-end of the primary reaction product.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 4 minutes and 30 seconds of 94°C followed by 25 cycles of 94°C, 30 seconds; 60°C, 45 seconds; 72°C, 1 minute, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from

the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(c). Third amplification. One µl (500 pg) of 100 times diluted gel-purified primary reaction product was re-amplified using 12.5pmol of each of the primers NNBREPFOR (SEQ ID 22) and ORIREV408 (SEQ ID 20) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The NNBREPFOR primer anneals to the 5'-end of the primary reaction product and appends a random amino acid 12-mer NNB library DNA. The ORIREV408 primer anneals to the 3'-end of the primary reaction product.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 4 minutes and 30 seconds of 94°C followed by 25 cycles of 94°C, 30 seconds; 60°C, 45 seconds; 72°C, 1 minute, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(d). Fourth amplification. One µl (500 pg) of 100 times diluted pTACP2A plasmid (ref) using 12.5pmol of each of the primers TACFARUP (SEQ ID 23) and TACREV (SEQ ID 27) in a 50µl reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TACFARUP primer anneals to the 5'-end of the TAC promoter DNA. The TACREV primer anneals to the 3'-end of the TAC promoter DNA.

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 1 minutes and 45 seconds of 94°C followed by 25 cycles of 94°C, 15 seconds; 60°C, 30 seconds; 72°C, 30 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40µl sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(e). First assembly PCR. One  $\mu$ l (50 ng) of each of the reaction products in (b) and (d) using 50 pmol of each of the primers TACFARUP (SEQ ID 23) and ORIREV408 (SEQ ID 20) in a 50 $\mu$ l reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TACFARUP primer anneals to the 5'-end of the reaction product (d). The ORIREV480 primer anneals to the 3'-end of the reaction product (b). The reaction product in (e) is called TAC-V5-REPA-CIS-ORI-408 (SEQ ID 24).

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 1 minutes and 45 seconds of 94°C followed by 25 cycles of 94°C, 15 seconds; 60°C, 30 seconds; 72°C, 1 minute and 30 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40 $\mu$ l sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(f). Second assembly PCR. One  $\mu$ l (50 ng) of each of the reaction products in (c) and (d) using 50 pmol of each of the primers TACFARUP (SEQ ID 23) and ORIREV408 (SEQ ID 20) in a 50 $\mu$ l reaction containing 0.25mM dNTPs, 2.5 units TaqDeepVent DNA polymerase mixture (20:1), and 1x PCR reaction buffer (New England Biolabs, Beverly, MA, U.S.A.). The TACFARUP primer anneals to the 5'-end of the reaction product (d). The ORIREV480 primer anneals to the 3'-end of the reaction product (c).

PCR reactions were carried out on a Eppendorf Master Cycler for 1 cycle of 1 minutes and 45 seconds of 94°C followed by 25 cycles of 94°C, 15 seconds; 60°C, 30 seconds; 72°C, 1 minute and 30 seconds, followed by a single cycle 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel into 40 $\mu$ l sterile water using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.). The reaction product in (f) is called TAC-NNB-REPA-CIS-ORI-408 (SEQ ID 25).

Preparation of *in vitro* transcription/translation reaction: The reaction set was set up on ice, using a Promega bacterial linear template S30 coupled *in vitro* transcription/translation reaction kit as follows:

20µl of 5000 times diluted TAC-V5-REPA-CIS-ORI-408 template (0.1ng of final construct DNASEQ ID 24 above)

20µl of 5 TAC-NNB-REPA-CIS-ORI-408 template (0.5µg of final construct DNASEQ ID 25 above)

20µl complete amino acid mix (Promega)

80µl S30 Premix

60µl S30 mix

and the reaction was allowed to proceed at 25°C for 30 minutes and placed on ice, then diluted 10 fold with 2% Marvel/PBS.

DNA-protein complex capture. NUNC star immunotubes were coated with 10µg/ml of anti-V5 antibody in 500µl PBS overnight at 4°C. An additional tube was left blank as a negative control. Tubes were washed 2x PBS and blocked for 1 hour at room temperature with blocking buffer (2% Marvel, 0.1% Tween 20, 0.1mg/ml herring sperm DNA) and then washed 2x PBS. 1 ml of diluted transcription/translation reaction was added to each tube and incubated at room temperature for 1 hour. Tubes were washed 5x PBS/0.1% Tween 20 and then 5x PBS. DNA was recovered with 500µl TE buffer plus 500µl phenol/chloroform. This was centrifuged at 13,200g for 5 minutes and DNA precipitated with 1/10 volume of 3M sodium acetate, 50µg/ml glycogen and two volumes of absolute ethanol. Following centrifugation, pellets were washed with 70% ethanol, vacuum dried and resuspended in 40µl water. 20 µl of recovered DNA was reamplified in 50 µl reactions with the biotinylated primers bTAC6 (SEQ ID 26) and bCISREV (SEQ ID 19). Reaction products were electrophoresed on a 1% agarose/TAE gel.

Cloning of recovered DNA into the expression vector pDMG-K (SEQ ID 27). Reaction product were gelpurified and eluted with 50µl sterile water using a QIAquick Gel extraction kit according to the manufacturers instructions (QIAGEN Ltd West Sussex, U.K.). Both the purified reaction product and the plasmid pDMG-K were digested with 20 units of NcoI and NotI (New England Biolabs, Beverly, MA,

U.S.A.). The cut plasmid was gelpurified using a QIAquick Gel extraction kit according to the manufacturers instructions (QIAGEN Ltd West Sussex, U.K.), then treated with 0.01 units of Calf Intestinal Alkaline Phosphatase (Promega, Southampton, U.K.) followed by phenol/chloroform extraction and ethanol precipitation as described above. Precipitated DNA was dissolved in 20µl of water. The cut PCR product was transferred to Streptavidin coated strips (Roche Diagnostics Ltd, East Sussex, U.K.) in 1x TBS, 0.3 mg/ml BSA, 0.1% Tween 20 and incubated for 30 minutes at room temperature, shaking. This approach removes the flanking biotinylated DNA upstream and downstream of the NcoI and NotI site of the PCR product and enables recovery of the small DNA fragment containing the selected peptide sequence. Supernatant was phenol/chloroform extracted and ethanol precipitated as described above. Precipitated DNA was dissolved in 10µl of water. Cut plasmid and the isolated small DNA fragment containing the selected peptide sequence, both having NcoI and NotI overhangs, were ligated using a Quick ligation kit according to the manufacturers instructions (New England Biolabs, Beverly, MA, U.S.A.) followed by phenol/chloroform extraction and ethanol precipitation as described above. Precipitated DNA was dissolved in 10µl of water and electroporated into electrocompetent TG1 cells according to the manufacturers instructions (Stratagene, U.S.A.) and selected on plates with 2xTY, 100µg/ml ampicillin, and 2% glucose.

Anti-V5 antibody ELISA screening of selected clones. 88 colonies were picked into 400µl of 2x TY, 2% glucose, and 100µg/ml ampicillin and grown overnight at 37°C, shaking 300 rpm. 50µl of the overnight cultures were transferred into 1ml of 2x TY, 2% glucose, and 100µg/ml ampicillin and grown at 37°C, shaking 300 rpm until OD 0.5. Then the cells were centrifuged at 1000x g for 10 minutes. The supernatants were discarded and pellets were resuspended in 600µl of 2x TY, 0.4M sucrose, 100µg/ml ampicillin, and 1mM IPTG and grown for 4 hours at 37°C, 300 rpm. After induction the cells were centrifuged at 1000x g for 10 minutes. 150µl of the supernatants were used in the ELISA test. NUNC Maxisorp plates were coated with 100µl of 1µg/ml in 1x PBS of either anti-human kappa region antibody or anti-V5 antibody or 50µg/ml of BSA for 7 hours at room temperature. An additional plate

was left blank, only coated with PBS. Wells were rinsed 2x PBS followed by blocking for 1 hour at room temperature with 300µl of 4% Marvel, 0.1% Tween in 1x PBS. Wells were rinsed 2x PBS, then 150µl of supernatant and 150µl of 4% Marvel, 0.1% Tween 20 in 1x PBS were added to wells and incubated for 1 hour at room temperature. Wells were then washed 2xPBS, 0.1% Tween 20 and 2x PBS. Secondary antibody anti-human kappa region antibody conjugated to horse radish peroxidase (HRP)(final concentration 1.6µg/ml) was diluted 500 times in 4% Marvel, 0.1% Tween 20, 1x PBS and added to wells and incubated for 1 hour at room temperature. Wells were then washed 4x PBS, 0.1% Tween 20 and 2x PBS. The HRP signal was detected by adding 200µl of TMB substrate. Reaction was stopped with 100µl of 0.5M sulphuric acid. Absorbance was read at 450nm. 35 out of 88 clones expressed well judged by HRP signal from clones screened against anti-human kappa region antibody. 7 out of these 35 clones showed specific binding to anti-V5 antibody, thereby enriching V5-peptides from 1 in 5000 to 1 in 5, i.e. an enrichment factor of 1000 (Figure 5).

**Example 4: CIS display Library construction, selection & screening against *Bacillus globigii***

**Library Construction**

To generate library DNA, a promoter library DNA fragment and the RepA-CIS-ori fragment must be generated, then linked together by digestion-ligation. The tac promoter from a P2A-HA vector was used in this example, but many available promoters could be used, and are well known to those skilled in the art. The initial PCR of Rep-CIS-ori and TAC fragments appends Bsp120I site and the random library/NotI site respectively. Two master mixes were prepared:

10µl of 1:50 diluted P2A-HA plasmid DNA (25ng/reaction) was PCR amplified in 20x 50µl reaction volume containing 200µM dNTPs, 1xNEB polymerase amplification buffer (10mM KCl, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20mM Tris-HCl pH 8.8, 2mM MgSO<sub>4</sub>, 0.1% TritonX-100) with 10pmol of each of the primers TACFARUP (SEQ ID 23) and NTERM18MER (SEQ ID 28) primers and 2 units of

20:1 Taq DNA polymerase: Deep Vent DNA polymerase mixture (NEB) for 25 cycles of 94°C, 40 seconds; 60°C, 40 seconds; 72°C, 60 seconds; followed by a 5 minutes extension at 72°C. 20µl of reaction product were electrophoresed on a 1% agarose/TAE gel and photographed, while the remainder was Qiagen column purified into 200µl water.

10µl of Bsp120I corrected Rep-CIS-ori DNA (50ng/reaction) was PCR amplified in 10x 50µl reaction volume containing 200µM dNTPs, 1xNEB polymerase amplification buffer (10mM KCl, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20mM Tris-HCl pH 8.8, 2mM MgSO<sub>4</sub>, 0.1% TritonX-100) with 10pmol of each of the primers BSPREPAFOR (SEQ ID 29) and ORIREV (SEQ ID 02) primers and 2 units of 20:1 Taq DNA polymerase: Deep Vent DNA polymerase mixture (NEB) for 30 cycles of 94°C, 40 seconds; 60°C, 40 seconds; 72°C, 90 seconds; followed by a 5 minutes extension at 72°C. 20µl of reaction product were electrophoresed on a 1% agarose/TAE gel and photographed, while the remainder was Qiagen column purified into 120µl water.

Library-TAC product was then digested with 10µl NotI (NEB) (100u) for 1 hour at 37°C in a 300µl reaction volume, then Qiagen column purified into a 120µl volume of water. The two products were then joined by restriction-ligation as follows:

10xNEB buffer 4	17µl
100mM ATP (SIGMA)	15µl
10mg/ml acetylated BSA (NEB)	1µl
RepA DNA	40µl
TAC-library DNA	40µl
Bsp120I (10u/µl Fermentas)	5µl
NotI (10u/µl NEB)	5µl
T4 DNA ligase (400u/µl NEB)	5µl
Water	39µl

Reaction was carried out at 37°C for two hours. 20µl was assessed by gel electrophoresis, 30µl was PCR amplified directly in 10x 50µl reactions, and the remainder was gel purified and the library band excised, Qiagen column purified and PCR amplified in 20x 50µl reactions with primers TACFAR4 (SEQ ID 30) and ORIREV (SEQ ID 02) for 30 cycles of 94°C, 40 seconds; 60°C, 40 seconds; 72°C, 90 seconds; followed by a 5 minutes extension at 72°C. DNA was gel purified in 4 Qiagen columns and the 200µl eluate pooled for ITT reactions/selection.

#### Round 1 Selection

2 x 200µl ITT reaction was set up and incubated at room temperature for 1 hour as follows:

REACTION	1
Library DNA	56µl (7µg)
2.5x buffer	80µl
100mM methionine	2µl
S30 extract	60µl

1ml of blocking buffer was added to each reaction (Block buffer is 4% Marvel, 100µg/ml sheared salmon sperm DNA, 0.1% Tween 20, 2.5mg/ml heparin, in TBS), spun at 10,000g for 2 minutes, transferred to a fresh tube, then placed on ice.

100 µl *Bacillus globigii* (Bg) spore suspension was washed twice with 1ml TBS/0.1% Tween 20 and was resuspended in 100µl of Block buffer. This was then added to the Block buffer and allowed to bind at room temperature for 1 hour whilst mixing.

The mix was then centrifuged at 16,100g for 1 minute and the spore pellet was washed six times with 1 ml of TBS/0.1% Tween 20 by mixing with a pipette and vortexing prior to centrifugation. The pellet was finally washed in 1ml TBS and the supernatant was discarded.



DNA was eluted by incubation of the spores in 120µl 0.5M sodium acetate pH5.5 for 10 minutes on a mixer. The spores were centrifuged at 16,100g for 1 minute and the supernatant was neutralised by the addition of 120µl Tris pH8.0 and then phenol/CHCl<sub>3</sub> extracted for 5 minutes at 16,100g. DNA was precipitated with 20µg carrier glycogen and two and a half volumes of ethanol. DNA was pelleted at 16,100g for 20 minutes and the pellet washed three times with 0.75ml 70% ethanol, centrifuging for 3 minutes at 16,100g in between each wash, then air dried and re-suspended in 20µl water.

10µl recovered DNA was PCR amplified in 10 × 50µl reaction with primers CISREV (SEQ ID 19) and TACFAR5 (SEQ ID 31) and 2 units of 20:1 Taq DNA polymerase: Deep Vent DNA polymerase mixture (NEB) for 30 cycles of 94°C, 40 seconds; 60°C, 40 seconds; 72°C, 90 seconds; followed by a 5 minutes extension at 72°C. The DNA was purified, ethanol precipitated and re-suspended in 10µl water. 5 µl were further amplified by PCR using the conditions above but using the primers NOTRECREV2 (SEQ ID 32) and TACFAR5 (SEQ ID 31) for 10 cycles. The product was purified using a Qiagen PCR purification kit and eluted into 50µl 5mM Tris pH 8.0.

#### Restriction-Ligation

This was carried out in a 30µl reaction for 1 hour at 37°C to reattach RepA-CIS-ori DNA to recovered peptides for a further round of selection.

---

10xNEB buffer 4	3µl
100mM ATP (SIGMA)	1.5µl
10mg/ml acetylated BSA (NEB)	0.3µl
RepA DNA	2µl
TAC-library DNA	10µl
Bsp120I (10u/µl Fermentas)	1.5µl
NotI (10u/µl NEB)	1.5µl
T4 DNA ligase (400u/µl NEB)	1.5µl

---

Water

8.7 $\mu$ l

---

20 $\mu$ l was PCR amplified directly in 10x 50 $\mu$ l reactions with primers TACFAR5.1 (SEQ ID 33) and ORIREV (SEQ ID 02) for 20 cycles of 94°C, 40 seconds; 60°C, 40 seconds; 72°C, 90 seconds; followed by a 5 minutes extension at 72°C. DNA was gel purified in 1 Qiagen column and the eluate used for Round 2 of ITT reactions/selection (58 $\mu$ l used in R2).

#### Round 2

Second round selection was carried out as for round 1, with the following changes: Approximately 3 $\mu$ g of input DNA were used. Block buffer was 2% bovine serum albumin, 1% gelatin, 100 $\mu$ g/ml sheared salmon sperm DNA, 2.5mg/ml heparin, in TBS. 10 $\mu$ l washed spores used in each selection. Recovery PCRs used TACFAR5.2 (SEQ ID 34) and NOTRECREV2 (SEQ ID 32) primers. Finally, pull through PCR used TACFAR5.2 (SEQ ID 34) and ORIREV (SEQ ID 02) primers for 10 cycles.

#### Round 3

Third round selection was carried out as for round 2, with the following changes: Approximately 2.5 $\mu$ g of input DNA was used. Recovery PCRs used TACFAR6 (SEQ ID 35) and NOTRECREV2 (SEQ ID 32) primers. Finally, pull through PCR used TACFAR6 (SEQ ID 35) and ORIREV (SEQ ID 02) primers for 10 cycles.

#### Round 4

Round 4 was carried out as for round 3, except that approximately 2 $\mu$ g of input DNA was used for the selection. Recovery PCRs used TAC3 (SEQ ID 09) and NOTRECREV2 (SEQ ID 32) primers. Finally, pull through PCR used TAC3 (SEQ ID 09) and ORIREV (SEQ ID 02) primers for 10 cycles.

### Round 5

Round 5 was carried out as for round 4.

For cloning out as NcoI-NotI fragments, the stored the recovered DNA from round 5 was PCR amplified with biotinylated TAC6 (SEQ ID 26) primer and NOTRECREV2 (SEQ ID 32). Digestion with NotI was followed by purification using Qiagen PCR purification kit, digestion with NcoI followed by incubation in a plate coated with streptavidin. Following phenol/CHCl<sub>3</sub> purification and ethanol precipitation, the digested DNA was then ligated into a similarly digested pVIII phagemid vector and transformed into ER2738 *E.coli*, then plated on 2% glucose, 2xTY, 100µg/ml ampicillin plates and incubated o/n at 37°C.

Individual colonies were picked into 200µl 2% glucose, 2xTY, 100µg/ml ampicillin medium in 96 well plates, and grown at 37°C/200rpm for 6 hours. 100µl was transferred to a deep-well plate containing 100µl 2% glucose, 2xTY, 100µg/ml ampicillin plus 10µl M13K07 helper phage/well and incubated for 1 hour without shaking at 37°C. 500µl per well of 2xTY, 100µg/ml ampicillin/ 25µg/ml kanamycin/ 20µM IPTG medium was added and incubation carried out o/n at 37°C/200rpm.

### ELISA screening

Round bottom 96 well plates were blocked with 4% Marvel in TBS/0.1%Tween 20 in PBS for 1 hour at room temperature. Picked phage cultures were centrifuged at 3000g for 5 minutes and the phage supernatant was assayed in an ELISA. In each well 50µl of phage supernatant were mixed with 5µl *Bg* spores in 50µl 4% bovine serum albumin, 1% gelatin in TBS and incubated whilst shaking at room temperature for 1 hour. The wells were washed 5 × with 200µl TBS/0.1% Tween20 by centrifugation at 3000g for 5 minutes in between each wash before incubation with anti-M13 horseradish peroxidase conjugated antibody 0.2µg/ml in 4% bovine serum albumin, 1% gelatin in TBS. The spores were incubated at room temperature for 1 hour whilst shaking. The wells were then washed 5 × with TBS/0.1% Tween20 and the spores were transferred into a fresh plate. The spores were then washed once with TBS as described above before development with TMB

substrate. The development was stopped with 0.5M H<sub>2</sub>SO<sub>4</sub> and the solution was transferred to a fresh flat-bottomed plate for reading at 450nm. Binding data for selected peptides is shown in Figure 6.

**Example 5. CIS display Library construction, selection & screening against anti-V5 antibody**

Library Construction was carried out as described in Example 4.

**Round 1 Selection**

1x 200µl in vitro transcription/translation reaction (ITT) reaction was set up and incubated at room temperature for 1 hour as described in Example 4. 1ml of blocking buffer was added to each reaction (Block buffer is 5% skimmed milk powder, 100µg/ml sheared salmon sperm DNA, 2.5mg/ml heparin, in TBS), then placed on ice.

For the first round of library selection a 70x11mm NUNC Maxisorp Immuntube (Life Technologies, Paisley, Scotland U.K.) was coated with 1ml of 10µg/ml of polyclonal anti-V5 peptide antibody (Harlan-Seralab) in PBS for 1 hour at 37°C. The tube was rinsed three times with PBS (fill & empty) and blocked with 3ml block buffer for 1 hour at 37°C and washed as before. Library protein-DNA complexes in block buffer were added, and incubated for 1 hour standing at room temperature. The tube was washed five times with PBS/0.1% Tween 20, then a further five times with PBS only.

DNA was eluted into 500µl 1M Sodium acetate pH 5.2 for 10 minutes on the blood mixer, neutralized with 100µl 1M Tris-HCl pH 8.0, then phenol/CHCl<sub>3</sub> extracted for 5 minutes at 16,100g. DNA was precipitated with 20µg carrier glycogen, ½ volume 7.5M ammonium acetate, and three volumes of ethanol. DNA was pelleted at 16,100g for 20 minutes and the pellet washed with 0.5ml 70% ethanol for 5 minutes at 16,100g then vacuum dried, and re-suspended in 20µl water.

10µl recovered DNA was PCR amplified in 1x 50µl reaction with primers NOT1RECREV2 (SEQ ID 32) and TACFAR4 (SEQ ID 30) and 2 units of 20:1 Taq DNA polymerase: Deep Vent DNA polymerase mixture (NEB) for 30 cycles of

94°C, 40 seconds; 60°C, 40 seconds; 72°C, 90 seconds; followed by a 5 minutes extension at 72°C. 50µl of reaction product were electrophoresed on a 1% agarose/TAE gel and photographed, then GeneClean purified into 10µl water. DNA was reattached to RepA DNA and reamplified for round two as described in example 4 using TACFAR5 (SEQ ID 31) and ORIREV (SEQ ID 02) primers.

Second round selection was carried out as for round 1, using the same primer pairs as described in example 4, with the following changes: Anti-V5 antibody coating concentration was reduced to 5µg/ml. Input DNA was approximately 4µg. Third round selection was carried out as for round 2, with the following changes: Approximately 4µg of input DNA was used. Recovery PCRs used TACFAR5.1 (SEQ ID 33) and NOTRECREV2 (SEQ ID 32) primers. Finally, pull through PCR used TACFAR6 (SEQ ID 35) and ORIREV (SEQ ID 02) primers for 10 cycles. Round 4 was carried out as for round 3.

For cloning out as NcoI-NotI fragments, the stored the recovered DNA from round 4 the recovered DNA from round 4 was PCR amplified with biotinylated (SEQ ID 26) TAC6 and NOTIREPRECREV2 (SEQ ID 32) primers and cloned into pVIII phagemid vector and electroporated into electrocompetent TG-1 *E.coli*, as described in example 4.

Individual colonies were picked into 200µl 2% glucose, 2xTY, 100µg/ml ampicillin medium in 96 well plates, and grown at 37°C/200rpm for 6 hours. 100µl was transferred to a deep-well plate containing 100µl 2% glucose, 2xTY, 100µg/ml ampicillin plus 10<sup>9</sup> kru M13K07 helper phage/well and incubated for 1 hour without shaking at 37°C. 400µl per well of 2xTY, 100µg/ml ampicillin/ 25µg/ml kanamycin/ 20µM IPTG medium was added and phage amplification continued for 16 hours at 37°C while shaking at 200rpm. Bacterial cultures were spun in microtitre plate carriers at 2000g for 10 minutes at 4°C in a benchtop centrifuge to pellet bacteria and culture supernatant used for ELISA.

A NUNC Maxisorp ELISA plate was coated with 100ng/well anti-V5 peptide antibody in 100µl /well PBS for one hour at 37°C. The plate was washed 2x200µl/well PBS and blocked for 1 hour at 37°C with 200µl/well 2% BSA/PBS and

then washed 2x200µl/well PBS. 50µl phage culture supernatant was added to each well containing 50µl/well 4% BSA/PBS, and allowed to bind for 1 hour at room temperature. The plate was washed two times with 200µl/well PBS/0.1% Tween 20, then two times with 200µl/well PBS. Bound phage were detected with 100µl/well, 1:5000 diluted anti-M13-HRP conjugate (Amersham-Pharmacia) in 2% BSA/PBS for 1 hour at room temperature and the plate washed four times as above. The plate was developed for 5 minutes at room temperature with 100µl/well TMB (3,3',5,5'-Tetramethylbenzidine) substrate buffer. The reaction was stopped with 100µl/well 0.5N H<sub>2</sub>SO<sub>4</sub> and read at 450nm. Phagemid DNA of ELISA positive clones were then sequenced with standard pUC forward and reverse sequencing primers. The amino acid sequence of these clones isolated is shown below. Four ELISA positive clones were grown in 10ml culture volumes and phage particles precipitated with PEG-NaCl and re-suspended in 1ml PBS and 50µl retested in ELISA as described above. OD450nm signals against anti-V5 and control anti-ACTH peptide antibody are shown in Figure 7.

Peptide sequences isolated after selection:

<b>P1C12 (SEQ ID 36)</b>	<b>C G C P T M A A R V R P V L N S K H</b>
<b>P2H1 (SEQ ID 37)</b>	<b>M T T V P V L M I S V</b>
<b>P1B5 (SEQ ID 38)</b>	<b>T L S T R H H N V I D R F N L R N F</b>
<b>P2B8 (SEQ ID 39)</b>	<b>S I R T L T G S T P A Q F D A T A D</b>

## SEQUENCE LISTING

<110> ISOGENICA LIMITED

---

<120> PEPTIDE LIBRARY DISPLAY METHOD

<130> P.86234A SER

<160> 39

<170> PatentIn version 3.1

<210> 1

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 1

actgatcttc accaaacgta tta

23

<210> 2

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 2

tgcatatctg tctgtccaca gg

22

<210> 3

<211> 48

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 3

gagcttcaac aggggagggg gaggaggatc aactgatctt caccaaac

48

<210> 4

<211> 50

---

<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 4  
ctaggactgg attcaacggg gggaggagga tcaactgac ttcaccaaac 50

<210> 5  
<211> 49  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 5  
cagaagagga tctgaatggg ggaggagggt ccactgtggc tgcaccatc 49

<210> 6  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 6  
tcccctgttg aagctctttg tg 22

<210> 7  
<211> 40  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 7  
cagaagagga tctgaatggg ggaggagggt ccggaacc 40

<210> 8  
<211> 27  
<212> DNA



<213> Artificial sequence

<220>

<223> Primer

<400> 8

gctacgttga atccagtcct aggagag

27

<210> 9

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 9

catattgtcg ttagaacgcg gc

22

<210> 10

<211> 58

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 10

attcagatcc tcttctgaga tgagtttttg ttcctcgagc atggtagatc ctgtttcc

58

<210> 11

<211> 42

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 11

cgatacctag cgttcggatc catattgtcg ttagaacgcg gc

42

<210> 12

<211> 1788

<212> DNA

<213> Artificial sequence

&lt;220&gt;

&lt;223&gt; DNA construct

&lt;400&gt; 12

catattgtcg	ttagaacgcg	gctacaatta	atacataacc	ttatgtatca	tacacatacg	60
athtaggtga	cactatagaa	tacaagctta	ctccccatcc	ccctgttgac	aattaatcat	120
ggctcgtata	atgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acaggatcta	180
ccatgctcga	ggaacaaaaa	ctcatctcag	aagaggatct	gaatggggga	ggagggtcca	240
ctgtggctgc	accatctgtc	ttcatcttcc	cgccatctga	tgagcagttg	aaatctggaa	300
ctgcctctgt	tgtgtgcctg	ctgaataact	tctatcccag	agaggccaaa	gtacagtggg	360
aggtggataa	cgccctccaa	tcgggtaact	cccaggagag	tgacacagag	caggacagca	420
aggacagcac	ctacagcctc	agcaacaccc	tgacgctgag	caaagcagac	tacgagaaac	480
acaaagtcta	cgcctgcgaa	gtcacccatc	agggcctgag	ctcgcccgtc	acaaagagct	540
tcaacagggg	agggggagga	ggatcaactg	atcttcacca	aacgtattac	cgccaggtaa	600
agaaccgcaa	tccggtgttc	actccccgtg	aaggtgccgg	aacgccgaag	ttccgcgaaa	660
aaccgatgga	aaaggcgggt	ggcctcacct	cccgttttga	tttcgccatt	catgtggcgc	720
atgcccggtc	ccgtgggtctg	cgctgggcga	tgccaccggt	gctgcgtcga	cgggctattg	780
atgcgctgct	gcaggggctg	tgtttccact	atgaccgctg	ggccaaccgc	gtccagtgtt	840
ccatcaccac	actggccatt	gagtgcggac	tggcgacaga	gtccggtgca	ggaaaactct	900
ccatcaccgg	tgccaccggg	gccctgacgt	tctgtcaga	gctgggactg	attacctacc	960
agacggaata	tgaccgcgtt	atcgggtgct	acattccgac	cgacatcacg	ttcacactgg	1020
ctctgtttgc	tgcccttgat	gtgtctgagg	atgcagtggc	agctgcgcgc	cgcagtcgtg	1080
ttgaatggga	aaacaaacag	cgcaaaaagc	aggggctgga	taccctgggt	atggatgagc	1140
tgatagcgaa	agcctggcgt	tttgtgcgtg	agcgtttccg	cagttaccag	acagagcttc	1200
agtcccgtgg	aataaaaacgt	gcccgtgcgc	gtcgtgatgc	gaacagagaa	cgtcaggata	1260
tcgtcacctc	agtgaaacgg	cagctgacgc	gtgaaatctc	ggaaggacgc	ttcactgcta	1320
atggtgaggc	ggtaaaacgc	gaagtggagc	gtcgtgtgaa	ggagcgcag	attctgtcac	1380
gtaaccgcaa	ttacagccgg	ctggccacag	cttctcccctg	aaagtgatct	cctcagaata	1440
atccggcctg	cgccggaggc	atccgcacgc	ctgaagcccc	ccggtgcaca	aaaaaacagc	1500
gtcgcagtcg	aaaaacaatc	tcatcatcca	ccttctggag	catccgattc	cccctgtttt	1560
taatacaaaa	tacgcctcag	cgacggggaa	ttttgcttat	ccacatttaa	ctgcaagggg	1620
cttccccata	aggttacaac	cgttcatgtc	ataaagcgcc	agccgccagt	cttacagggt	1680
gcaatgtatc	ttttaaacac	ctgtttatat	ctcctttaaa	ctacttaatt	acattcattt	1740
aaaaagaaaa	cctatttact	gcctgtcctg	tggacagaca	gatatgca		1788

&lt;210&gt; 13

&lt;211&gt; 1518

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; DNA construct

&lt;400&gt; 13

catattgtcg	ttagaacgcg	gctacaatta	atacataacc	ttatgtatca	tacacatacg	60
athtaggtga	cactatagaa	tacaagctta	ctccccatcc	ccctgttgac	aattaatcat	120
ggctcgtata	atgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acaggatcta	180
ccatgctcga	ggaacaaaaa	ctcatctcag	aagaggatct	gaatggggga	ggagggtccg	240
gaaaacctat	cccaaaccct	ctcctaggac	tggattcaac	ggggggagga	ggatcaactg	300

atcttcacca	aacgtattac	cgccaggtaa	agaacccgaa	tccggtgttc	actccccgtg	360
aaggtgccgg	aacgccgaag	ttccgcgaaa	aaccgatgga	aaaggcgggtg	ggcctcacct	420
cccgttttga	tttcgccatt	catgtggcgc	atgcccgttc	ccgtgggtctg	cgtcggcgca	480
<del>tgecaecgg</del>	<del>getgegtcga</del>	<del>cgggctattg</del>	<del>atgcgctgct</del>	<del>gcaggggctg</del>	<del>tgtttccact</del>	<del>540</del>
atgacccgct	ggccaaccgc	gtccagtgtt	ccatcaccac	actggccatt	gagtgcggac	600
tggcgacaga	gtccggtgca	ggaaaactct	ccatcaccgc	tgccaccg	gccctgacgt	660
tcctgtcaga	gctgggactg	attacctacc	agacggaata	tgaccgctt	atcgggtgct	720
acattccgac	cgacatcacg	ttcacactgg	ctctgtttgc	tgcccttgat	gtgtctgagg	780
atgcagtggc	agctgcgcgc	cgcagtcgtg	ttgaatggga	aaacaaacag	cgcaaaaagc	840
aggggctgga	taccctgggt	atggatgagc	tgatagcgaa	agcctggcgt	tttgtgcgtg	900
agcgtttccg	cagttaccag	acagagcttc	agtcccgtgg	aataaaacgt	gcccgtcgc	960
gtcgtgatgc	gaacagagaa	cgtcaggata	tcgtcaccct	agtgaacgg	cagctgacgc	1020
gtgaaatctc	ggaaggacgc	ttcactgcta	atgggtgaggc	ggtaaacgc	gaagtggagc	1080
gtcgtgtgaa	ggagcgcag	attctgtcac	gtaaccgcaa	ttacagccg	ctggccacag	1140
cttctccctg	aaagtgatct	cctcagaata	atccggcctg	cgccggaggc	atccgcacgc	1200
ctgaagcccg	ccggtgcaca	aaaaaacagc	gtcgcagca	aaaaacaatc	tcatcatcca	1260
ccttctggag	catccgattc	cccctgtttt	taatacaaaa	tacgcctcag	cgacggggaa	1320
ttttgcttat	ccacatttaa	ctgcaaggga	cttccccata	aggttacaac	cgttcatgtc	1380
ataaagcgcc	agccgccagt	cttacagggt	gcaatgtatc	ttttaaacac	ctgtttatat	1440
ctcctttaa	ctacttaatt	acattcattt	aaaaagaaaa	cctattcact	gcctgtcctg	1500
tggacagaca	gatatgca					1518

&lt;210&gt; 14

&lt;211&gt; 38

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Estrogen Receptor Target Recognition Sequence

&lt;400&gt; 14

tcaggtcaga gtagcctgag ctaaaataac acattcag

38

&lt;210&gt; 15

&lt;211&gt; 828

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; repA sequence

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(828)

&lt;223&gt;

&lt;400&gt; 15

atg gta aag aac ccg aat ccg gtg ttc act ccc cgt gaa ggt gcc gga

48

Met Val Lys Asn Pro Asn Pro Val Phe Thr Pro Arg Glu Gly Ala Gly	
1 5 10 15	
acg ccg aag ttc cgc gaa aaa ccg atg gaa aag gcg gtg ggc ctc acc	96
Thr Pro Lys Phe Arg Glu Lys Pro Met Glu Lys Ala Val Gly Leu Thr	
20 25 30	
tcc cgt ttt gat ttc gcc att cat gtg gcg cat gcc cgt tcc cgt ggt	144
Ser Arg Phe Asp Phe Ala Ile His Val Ala His Ala Arg Ser Arg Gly	
35 40 45	
ctg cgt cgg cgc atg cca ccg gtg ctg cgt cga cgg gct att gat gcg	192
Leu Arg Arg Arg Met Pro Pro Val Leu Arg Arg Arg Ala Ile Asp Ala	
50 55 60	
ctg ctg cag ggg ctg tgt ttc cac tat gac ccg ctg gcc aac cgc gtc	240
Leu Leu Gln Gly Leu Cys Phe His Tyr Asp Pro Leu Ala Asn Arg Val	
65 70 75 80	
cag tgt tcc atc acc aca ctg gcc att gag tgc gga ctg gcg aca gag	288
Gln Cys Ser Ile Thr Thr Leu Ala Ile Glu Cys Gly Leu Ala Thr Glu	
85 90 95	
tcc ggt gca gga aaa ctc tcc atc acc cgt gcc acc cgg gcc ctg acg	336
Ser Gly Ala Gly Lys Leu Ser Ile Thr Arg Ala Thr Arg Ala Leu Thr	
100 105 110	
ttc ctg tca gag ctg gga ctg att acc tac cag acg gaa tat gac ccg	384
Phe Leu Ser Glu Leu Gly Leu Ile Thr Tyr Gln Thr Glu Tyr Asp Pro	
115 120 125	
ctt atc ggg tgc tac att ccg acc gac atc acg ttc aca ctg gct ctg	432
Leu Ile Gly Cys Tyr Ile Pro Thr Asp Ile Thr Phe Thr Leu Ala Leu	
130 135 140	
ttt gct gcc ctt gat gtg tct gag gat gca gtg gca gct gcg cgc cgc	480
Phe Ala Ala Leu Asp Val Ser Glu Asp Ala Val Ala Ala Ala Arg Arg	
145 150 155 160	
agt cgt gtt gaa tgg gaa aac aaa cag cgc aaa aag cag ggg ctg gat	528
Ser Arg Val Glu Trp Glu Asn Lys Gln Arg Lys Lys Gln Gly Leu Asp	
165 170 175	
acc ctg ggt atg gat gag ctg ata gcg aaa gcc tgg cgt ttt gtg cgt	576
Thr Leu Gly Met Asp Glu Leu Ile Ala Lys Ala Trp Arg Phe Val Arg	
180 185 190	
gag cgt ttc cgc agt tac cag aca gag ctt cag tcc cgt gga ata aaa	624
Glu Arg Phe Arg Ser Tyr Gln Thr Glu Leu Gln Ser Arg Gly Ile Lys	
195 200 205	

cgt gcc cgt gcg cgt cgt gat gcg aac aga gaa cgt cag gat atc gtc 672  
 Arg Ala Arg Ala Arg Arg Asp Ala Asn Arg Glu Arg Gln Asp Ile Val  
 210 215 220

acc cta gtg aaa cgg cag ctg acg cgt gaa atc tcg gaa gga cgc ttc 720  
 Thr Leu Val Lys Arg Gln Leu Thr Arg Glu Ile Ser Glu Gly Arg Phe  
 225 230 235 240

act gct aat ggt gag gcg gta aaa cgc gaa gtg gag cgt cgt gtg aag 768  
 Thr Ala Asn Gly Glu Ala Val Lys Arg Glu Val Glu Arg Arg Val Lys  
 245 250 255

gag cgc atg att ctg tca cgt aac cgc aat tac agc cgg ctg gcc aca 816  
 Glu Arg Met Ile Leu Ser Arg Asn Arg Asn Tyr Ser Arg Leu Ala Thr  
 260 265 270

gct tct ccc tga 828  
 Ala Ser Pro  
 275

<210> 16  
 <211> 275  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> repA sequence

<400> 16  
 Met Val Lys Asn Pro Asn Pro Val Phe Thr Pro Arg Glu Gly Ala Gly  
 1 5 10 15

Thr Pro Lys Phe Arg Glu Lys Pro Met Glu Lys Ala Val Gly Leu Thr  
 20 25 30

Ser Arg Phe Asp Phe Ala Ile His Val Ala His Ala Arg Ser Arg Gly  
 35 40 45

Leu Arg Arg Arg Met Pro Pro Val Leu Arg Arg Arg Ala Ile Asp Ala  
 50 55 60

Leu Leu Gln Gly Leu Cys Phe His Tyr Asp Pro Leu Ala Asn Arg Val  
 65 70 75 80

Gln Cys Ser Ile Thr Thr Leu Ala Ile Glu Cys Gly Leu Ala Thr Glu  
 85 90 95

Ser Gly Ala Gly Lys Leu Ser Ile Thr Arg Ala Thr Arg Ala Leu Thr  
                   100                  105                  110

Phe Leu Ser Glu Leu Gly Leu Ile Thr Tyr Gln Thr Glu Tyr Asp Pro  
           115                  120                  125

Leu Ile Gly Cys Tyr Ile Pro Thr Asp Ile Thr Phe Thr Leu Ala Leu  
       130                  135                  140

Phe Ala Ala Leu Asp Val Ser Glu Asp Ala Val Ala Ala Ala Arg Arg  
   145                  150                  155                  160

Ser Arg Val Glu Trp Glu Asn Lys Gln Arg Lys Lys Gln Gly Leu Asp  
                   165                  170                  175

Thr Leu Gly Met Asp Glu Leu Ile Ala Lys Ala Trp Arg Phe Val Arg  
           180                  185                  190

Glu Arg Phe Arg Ser Tyr Gln Thr Glu Leu Gln Ser Arg Gly Ile Lys  
       195                  200                  205

Arg Ala Arg Ala Arg Arg Asp Ala Asn Arg Glu Arg Gln Asp Ile Val  
       210                  215                  220

Thr Leu Val Lys Arg Gln Leu Thr Arg Glu Ile Ser Glu Gly Arg Phe  
   225                  230                  235                  240

Thr Ala Asn Gly Glu Ala Val Lys Arg Glu Val Glu Arg Arg Val Lys  
           245                  250                  255

Glu Arg Met Ile Leu Ser Arg Asn Arg Asn Tyr Ser Arg Leu Ala Thr  
           260                  265                  270

Ala Ser Pro  
       275

<210> 17  
 <211> 172  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> CIS DNA element

<400> 17  
 aagtgatctc ctcagaataa tccggcctgc gccggaggca tccgcacgcc tgaagcccgc 60  
 cggtgcacaa aaaaacagcg tcgcatgcaa aaaacaatct catcatccac cttctggagc 120  
 atccgattcc cctgttttt aatacaaat acgcctcagc gacggggaat tt 172

<210> 18  
 <211> 195

<212> DNA

<213> Artificial sequence

<220>

<223> ori sequence

<400> 18

tgcttatcca catttaactg caagggactt ccccataagg ttacaaccgt tcatgtcata	60
aagcgccagc cgccagtctt acaggggtgca atgtatcttt taaacacctg tttatatctc	120
ctttaaacta cttaattaca ttcatttaaa aagaaaacct attcactgcc tgcctgtgg	180
acagacagat atgca	195

<210> 19

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 19

aattccccgt cgctgaggcg	20
-----------------------	----

<210> 20

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 20

cgtaagccgg tactgattga	20
-----------------------	----

<210> 21

<211> 110

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 21

cacaggaaac aggatctacc atggccggaa aacctatccc aaaccctctc ctaggactgg	60
---	----

attcaacggg gggaggagga tcagcgccg caactgatct tcaccaaacg

110

<210> 22  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<220>  
<221> misc\_feature  
<222> (29)..(30)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (32)..(33)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (35)..(36)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (38)..(39)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (41)..(42)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (44)..(45)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (47)..(48)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (50)..(51)  
<223> n = a, g, c or t



<220>  
 <221> misc\_feature  
 <222> (53)..(54)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (56)..(57)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (59)..(60)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (62)..(63)  
 <223> n = a, g, c or t

<400> 22  
 cacacaggaa acaggatcta ccatggccnn bnnbnnbnnb nnnbnnbnnb nnnbnnbnnb 60  
 bnnbggggga ggaggatcag cgccgcaac tgatcttcac caaacg 106

<210> 23  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer

<400> 23  
 cagttgatcg gcgcgagatt 20

<210> 24  
 <211> 2390  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> TAC-V5-REPA-CIS-ORI-408 construct

<400> 24  
 cagttgatcg gcgcgagatt taatcgccgc gacaatttgc gacggcgcggt gcagggccag 60  
 actggagggtg gcaacgccaa tcagcaacga ctgtttgccc gccagttgtt gtgccacgcg 120  
 gttgggaatg taattcagct cgccatcgc cgcttccact ttttccgcg ttttcgcaga 180

aacgtggctg	gcctggttca	ccacgcggga	aacggtctga	taagagacac	cggcatactc	240
tgcgacatcg	tataacgtta	ctggtttcac	attcaccacc	ctgaattgac	tctcttccgg	300
gcgctatcat	gccataccgc	gaaaggtttt	gcaccattcg	gctagcgatg	accctgctga	360
ttggttcgct	gaccatttcc	ggggtgcgga	acggcggttac	cagaaactca	gaaggttcgt	420
ccaaccaaac	cgactctgac	ggcagtttac	gagagagatg	atagggctctg	cttcagtaag	480
ccagatgcta	cacaattagg	cttgatcata	ttgtcggttag	aacgcgggcta	caattaatac	540
ataaccttat	gtatcataca	catacgattt	aggtgacact	atagaatata	agcttactcc	600
ccatccccct	gttgacaatt	aatcatggct	cgtataatgt	gtggaattgt	gagcggataa	660
caatttcaca	caggaaacag	gatctaccat	ggccggaaaa	cctatcccaa	accctctcct	720
aggactggat	tcaacggggg	gaggaggatc	agcggccgca	actgatcttc	accaaacgta	780
ttaccgccag	gtaaaagaacc	cgaatccggt	gtttacaccc	cgtgaagggtg	caggaaacgct	840
gaagtctctg	gaaaaactga	tggaaaaggc	ggtgggcttc	acttcccgtt	ttgatttcgc	900
cattcatgtg	gcgcatgccc	gttcgcgtgg	tctgcgtcga	cgcatgccac	cagtgtctgcg	960
tcgacgggct	attgatgcgc	tcctgcaggg	gctgtgtttc	cactatgacc	cgctggccaa	1020
ccgcgtccag	tgctccatca	ccacgtggc	cattgagtgc	ggactggcga	cggagtctgc	1080
tgccggaaaa	ctctccatca	cccgtgccac	ccgggccctg	acgttcctgt	cagagctggg	1140
actgattacc	taccagacgg	aatatgaccc	gcttatcggg	tgctacattc	cgaccgatata	1200
cacgttcaca	tctgactgt	ttgtgccct	cgatgtatca	gaggaggcag	tggccgccgc	1260
gcgccgcagc	cgtgtggtat	gggaaaacaa	acaacgcaaa	aagcaggggc	tggataccct	1320
gggcattgat	gaactgatag	cgaaagcctg	gcgttttgtt	cgtgagcgtt	ttcgcagtta	1380
tcagacagag	cttaagtccc	ggggaataaa	gcgtgcccg	gcgcgtcgtg	atgcggacag	1440
ggaacgtcag	gatattgtca	ccctggtgaa	acggcagctg	acgcgcgaaa	tcgcggaaag	1500
gcgcttca	gccaatcgtg	aggcggtaaa	acgcgaagtt	gagcgtcgtg	tgaaggagcg	1560
catgattctg	tcacgtaacc	gtaattacag	ccggctggcc	acagcttccc	cctgaaagtg	1620
acctcctctg	aataatccgg	cctgcgccgg	aggcttccgc	acgtctgaag	cccgacagcg	1680
cacaaaaaat	cagcaccaca	tacaaaaaac	aacctcatca	tccagcttct	ggtgcatccg	1740
gccccccctg	ttttcgatac	aaaacacgcc	tcacagacgg	ggaattttgc	ttatccacat	1800
taaactgcaa	gggacttccc	cataagggtta	caaccgttca	tgctataaag	cgccatccgc	1860
cagcggtaca	gggtgcaatg	tatcttttaa	acacctgttt	atatctcctt	taaactactt	1920
aattacattc	atttaaaaag	aaaacctatt	cactgcctgt	cctgtggaca	gacagatatg	1980
cacctccac	cgaaagcggc	gggccctac	cggagccgct	ttagttacaa	cactcagaca	2040
caaccaccag	aaaaaccccc	gtccagcgca	gaactgaaac	cacaaagccc	ctccctcata	2100
actgaaaagc	ggccccgccc	cggcccgaag	ggccggaaca	gagtcgcttt	taattatgaa	2160
tggtgtaact	acttcatcat	cgctgtcagt	cttctcgtg	gaagttctca	gtacacgctc	2220
gtaagcggcc	ctgacggccc	gctaacgcgg	agatacgccc	cgacttcggg	taaaccctcg	2280
tcgggaccac	tccgaccgcg	cacagaagct	ctctcatggc	tgaaagcggg	tatggtctgg	2340
cagggtctgg	gatgggtaag	gtgaaatcta	tcaatcagta	ccggcttacg		2390

<210> 25  
 <211> 2384  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> TAC-NNB-REPA-CIS-ORI-408 construct

<220>  
 <221> misc\_feature  
 <222> (695)..(696)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (698)..(699)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (701)..(702)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (704)..(705)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (707)..(708)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (710)..(711)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (713)..(714)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (716)..(717)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (719)..(720)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (722)..(723)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (725)..(726)

<223> n = a, g, c or t

<220>

<221> misc\_feature

<222> (728)..(729)

<223> n = a, g, c or t

<400> 25

cagttgatcg	gcgcgagatt	taatcgccgc	gacaatttgc	gacggcgcgt	gcagggccag	60
actggaggtg	gcaacgcaa	tcagcaacga	ctgtttgcc	gccagttgtt	gtgccacgcg	120
gttggaatg	taattcagct	ccgccatcgc	cgcttccact	tttcccgcg	tttctcgaga	180
aacgtggctg	gcctggttca	ccacgcggga	aacggtctga	taagagacac	cggcatactc	240
tgcgacatcg	tataacgtta	ctggtttcac	attcaccacc	ctgaattgac	tctcttccgg	300
gcgctatcat	gccataccgc	gaaagggttt	gcaccattcg	gctagcgatg	accctgctga	360
ttggttcgct	gaccatttcc	ggggtgcgga	acggcgttac	cagaaactca	gaaggttcgt	420
ccaaccaaac	cgactctgac	ggcagtttac	gagagagatg	atagggtctg	cttcagtaag	480
ccagatgcta	cacaattagg	cttgtacata	ttgtcgttag	aacgcggcta	caattaatac	540
ataaccttat	gtatcataca	catacgattt	agtgacact	atagaataca	agcttactcc	600
ccatccccct	gttgacaatt	aatcatggct	cgtataatgt	gtggaattgt	gagcggataa	660
caatttcaca	caggaaacag	gatctaccat	ggccnnbnnb	nbnbnbnbn	nbnbnbnbn	720
bnnbnbnbn	gggggaggag	gatcagcggc	cgcaactgat	cttcaccaa	cgtattaccg	780
ccaggtaaag	aacccgaatc	cggtgtttac	accccgtaga	ggtgcaggaa	cgctgaagtt	840
ctgcgaaaaa	ctgatggaaa	aggcgggtgg	cttcacttcc	cgttttgatt	tgcctattca	900
tgtggcgcat	gcccgttcgc	gtggtctgcg	tcgaecgatg	ccaccagtgc	tgcgtcgacg	960
ggctattgat	gcgctcctgc	aggggctgtg	tttccactat	gacccgctgg	ccaaccgcgt	1020
ccagtgtctc	atcaccacgc	tggccattga	gtgcggactg	gcgacggagt	ctgctgccgg	1080
aaaactctcc	atcaccgcgt	ccaccggggc	cctgacgttc	ctgtcagagc	tgggactgat	1140
tacctaccag	acggaatatg	acccgcttat	cgggtgctac	attccgaccg	atatcacggt	1200
cacatctgca	ctgtttgctg	ccctcgatgt	atcagaggag	gcagtggccg	ccgcgcgccg	1260
cagccgtgtg	gtatgggaaa	acaaacaacg	caaaaagcag	gggctggata	ccctgggcat	1320
ggatgaactg	atagcgaaag	cctggcggtt	tgttcgtgag	cgttttcgca	gttatcagac	1380
agagcttaag	tcccggggaa	taaagcgtgc	ccgtgcgcgt	cgtgatgcgg	acagggaacg	1440
tcaggatatt	gtcaccctgg	tgaacggca	gctgacgcgc	gaaatcgcg	aagggcgctt	1500
cactgccaat	cgtgaggcgg	taaaacgcga	agttgagcgt	cgtgtgaagg	agcgcgatgat	1560
tctgtcacgt	aaccgtaatt	acagccggct	ggccacagct	tccccctgaa	agtacacctc	1620
tctgaataat	ccggcctgcg	ccggaggcct	ccgcacgtct	gaagcccgcg	agcgacacaa	1680
aaatcagcac	cacatacaaa	aaacaacctc	atcatccagc	ttctggtgca	tccggccccc	1740
cctgttttcg	atacaaaa	cgcttcacag	acggggaatt	ttgcttatcc	acattaaact	1800
gcaagggact	tcccataaag	gttacaaccg	ttcatgtcat	aaagcgccat	ccgccagcgt	1860
taacagggtgc	aatgtatctt	ttaaacacct	gtttatatct	cctttaaact	acttaattac	1920
attcatttaa	aaagaaaacc	tattcactgc	ctgtcctgtg	gacagacaga	tatgcacctc	1980
ccaccgcaag	cggcgggccc	ctaccggagc	cgcttttagt	acaacactca	gacacaacca	2040
ccagaaaaac	cccgtccag	cgcagaactg	aaaccacaaa	gcccctccct	cataactgaa	2100
aagcggcccc	gccccggccc	gaagggccgg	aacagagtcg	cttttaatta	tgaatgttgt	2160
aactacttca	tcatcgctgt	cagtcttctc	gctggaagtt	ctcagtacac	gctcgtaagc	2220
ggccctgacg	gcccgctaac	gcggagatac	gccccgactt	cgggtaaacc	ctcgtcggga	2280
ccactccgac	cgcgcacaga	agctctctca	tggctgaaag	cgggtatggt	ctggcagggc	2340
tggggatggg	taaggtgaaa	tctatcaatc	agtaccggct	tacg		2384

<210> 26  
<211> 26  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 26  
cccatcccc ctgttgacaa ttaatc

26

<210> 27  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 27  
ggtagatcct gtttcctgtg tg

22

<210> 28  
<211> 110  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<220>  
<221> misc\_feature  
<222> (37)..(38)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (40)..(41)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (43)..(44)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature

<222> (46)..(47)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (49)..(50)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (52)..(53)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (55)..(56)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (58)..(59)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (61)..(62)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (64)..(65)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (67)..(68)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (70)..(71)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature  
<222> (73)..(74)  
<223> n = a, g, c or t

<220>  
<221> misc\_feature

<222> (76)..(77)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (79)..(80)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (82)..(83)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (85)..(86)  
 <223> n = a, g, c or t

<220>  
 <221> misc\_feature  
 <222> (88)..(89)  
 <223> n = a, g, c or t

<400> 28  
 acataccgtc atgcggccgc tgatcctcct ccccvnnvn nvnnvnnvn vnnvnnvnnv 60  
 nnvnnvnnvn nvnnvnnvn vnnvnnvng gccatggtag atcctgtttc 110

<210> 29  
 <211> 49  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer

<400> 29  
 ctggagatgg catcaagggc cccaactgat cttcaccaaa cgtattacc 49

<210> 30  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer

<400> 30

ggcgctatca tgccataccg

20

<210> 31

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 31

accattcggc tagcgatgac

20

<210> 32

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 32

ggtgaagatc agttgcggcc gctgacctc ctc

33

<210> 33

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 33

gattggttcg ctgaccattt cc

22

<210> 34

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 34

cggcgttacc agaaactcag a

21



<210> 35  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer

<400> 35  
 aaccgactct gacggcagtt

20

<210> 36  
 <211> 18  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> Peptide

<400> 36  
 Cys Gly Cys Pro Thr Met Ala Ala Arg Val Arg Pro Val Leu Asn Ser  
 1 5 10 15

Lys His

<210> 37  
 <211> 11  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> Peptide

<400> 37  
 Met Thr Thr Val Pro Val Leu Met Ile Ser Val  
 1 5 10

<210> 38  
 <211> 18  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> Peptide

&lt;400&gt; 38

Thr	Leu	Ser	Thr	Arg	His	His	Asn	Val	Ile	Asp	Arg	Phe	Asn	Leu	Arg
1				5					10					15	

Asn Phe

&lt;210&gt; 39

&lt;211&gt; 18

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Peptide

&lt;400&gt; 39

Ser	Ile	Arg	Thr	Leu	Thr	Gly	Ser	Thr	Pro	Ala	Gln	Phe	Asp	Ala	Thr
1				5					10					15	

Ala Asp

CLAIMS

1. ~~A method for producing an *in-vitro* peptide expression library~~  
comprising a plurality of peptides, wherein each peptide is linked to the DNA construct encoding the peptide, comprising the steps of:
  - (a) providing a DNA construct comprising:
    - (i) a DNA target sequence;
    - (ii) DNA encoding a library member peptide; and
    - (iii) DNA encoding a peptide capable of non-covalently binding directly or indirectly to said DNA target sequence of (i);wherein said DNA construct and encoded protein are selected to have cis-activity
  - (b) expressing a plurality of DNA constructs according to (a) wherein said DNA constructs encode a plurality of library member peptides such that each expressed peptide is non-covalently linked to the DNA from which it was produced.
2. A method according to claim 1 wherein said DNA construct further comprises:
  - (iv) a cis-acting DNA element.
3. A method according to claim 2 wherein said DNA construct of (a) further comprises
  - (v) DNA encoding a fragment comprising at least the C-terminal 20 amino acids of a repA protein wherein said fragment is capable of interacting with said cis-acting DNA element of (iv);wherein said cis-acting DNA element of (iv) is located 3' to said DNA of (ii), (iii) and (v).

4. A method according to any one of the preceding claims wherein the peptide encoded by said DNA of (iii) is capable of recognising and directly binding said DNA target sequence of (i).

5. A method according to claim 4 wherein the peptide encoded by said DNA of (iii) is a repA protein and wherein said DNA target sequence of (i) is ori.

6. A method according to any one of claims 3 to 5 wherein said repA is selected from repA of the IncI complex plasmids and repA of the IncF, IncB, IncK, IncZ and IncL/M plasmids.

7. A method according to claim 5 wherein said DNA construct comprises the sequence encoding repA, the cis DNA element and the ori DNA of the IncFII plasmid R1.

8. A method according to any one of the preceding claims wherein said repA protein has the sequence given in SEQ ID NO: 16 and wherein said cis DNA element has the sequence given in SEQ ID NO: 17.

9. A method according to claim 4 wherein the peptide encoded by said DNA of (iii) is an oestrogen receptor DNA binding domain and wherein said DNA target sequence of (i) is an oestrogen receptor target sequence.

10. A method according to claim 8 wherein said DNA binding domain comprises amino acids 176 to 282 of the oestrogen receptor DNA binding fragment and wherein said DNA target sequence comprises the oestrogen receptor target sequence given in SEQ ID NO: 14.

11. A method according to claim 1, 2 or 3 wherein the peptide encoded by said DNA of (iii) indirectly binds said DNA target sequence of (i) via a bifunctional

agent, one part of which binds said DNA target sequence of (i) and a second part of which binds the peptide encoded by said DNA of (iii).

12. A method according to claim 11 wherein said DNA target sequence comprises a DNA tag capable of being bound by said bifunctional agent, said tag being optionally selected from biotin and fluorescein.

13. A method according to claim 11 or 12 wherein the binding activities of said bifunctional agent are conferred by means of two antibodies or fragments thereof.

14. A method according to claim 13 wherein one or both of said binding activities are conferred by means of an Fab fragment.

15. A method according to any one of claims 11 to 14 wherein said bifunctional agent is provided prior to step (b).

16. A method according to claim 11 wherein said bifunctional agent is bound to said DNA target sequence of (i) and is capable of binding to the peptide encoded by said DNA of (iii).

17. A method according to claim 16 wherein said bifunctional agent is a polymer.

18. A method according to any one of the preceding claims wherein said DNA is under the control of suitable promoter and translation sequences to allow for *in vitro* transcription and translation.

19. A method according to any one of the preceding claims wherein said library member peptide is an antibody or fragment thereof.

20. A method according to any one of the preceding claims wherein said library comprises at least  $10^4$  molecules.

21. A method according to any one of the preceding claims wherein said expression is carried out in a coupled bacterial transcription/translation environment.

22. A method according to claim 21 wherein said coupled bacterial transcription/translation environment is the S30 extract system.

23. A method for producing an *in vitro* peptide expression library comprising a plurality of peptides, wherein each peptide is linked to the DNA construct encoding the peptide, comprising the steps of:

- (a) providing a DNA construct comprising:
  - (i) DNA encoding a library member peptide; and
  - (ii) DNA encoding a peptide capable of binding to a bifunctional agent;wherein said DNA construct and encoded protein are selected to have cis-activity;
- (b) binding a bifunctional agent or a DNA tag capable of binding a bifunctional agent to said DNA construct of (a), wherein said bifunctional agent is capable of binding to the peptide encoded by said DNA of (ii); and
- (c) expressing a plurality of DNA constructs according to (b), wherein said DNA constructs encode a plurality of library member peptides such that each expressed peptide is linked via said bifunctional agent to the DNA from which it was produced.

24. A method of identifying and/or purifying a peptide exhibiting desired properties from an *in vitro* peptide expression library produced according to the method of any one of the preceding claims, comprising at least the steps of (a) screening said library and (b) selecting and isolating the relevant library member.

25. A method of identifying a specific ligand binding peptide, said method comprising at least the steps of (a) screening an in vitro peptide expression library produced according to the method of any one of claims 1 to 23 with ligand molecules which are optionally bound to a solid support; (b) selecting and isolating a library member binding to said ligand molecule; and (c) isolating the peptide which binds specifically to said ligand molecule.

26. A method according to claim 24 or 25 wherein said library member peptides are antibodies or fragments thereof.

27. A method of identifying and/or purifying a peptide having the ability to bind a specific DNA target sequence comprising at least the steps of

- (a) providing an in vitro expression library according to any one of claims 1 to 23 wherein the peptide encoded by the DNA of (iii) is a library member peptide having DNA binding activity and wherein said DNA target sequence of (i) is the target sequence of interest;
- (b) selecting and isolating a library member in which the encoded protein binds to said target sequence; and
- (c) isolating the peptide which binds to said target sequence.

28. A method according to claim 27 wherein said library member peptides are zinc finger proteins, helix-loop-helix proteins or helix-turn-helix proteins.

29. A method according to any one of claims 24 to 28 wherein additionally the DNA expressing said isolated peptide is isolated.

30. An in vitro peptide expression library produced according to the method of any one of claims 1 to 23.

31. A DNA construct as described in any one of claims 1 to 23.

DISPLAY LIBRARYAbstract

The invention provides a method for making *in vitro* peptide expression libraries, and for the isolation of nucleotide sequences encoding peptides of interest, wherein the peptides or proteins are specifically associated with the DNA encoding them through non-covalent protein:DNA binding. The method describes ways of making the library itself, DNA molecules encoding the library and uses of the expression library.



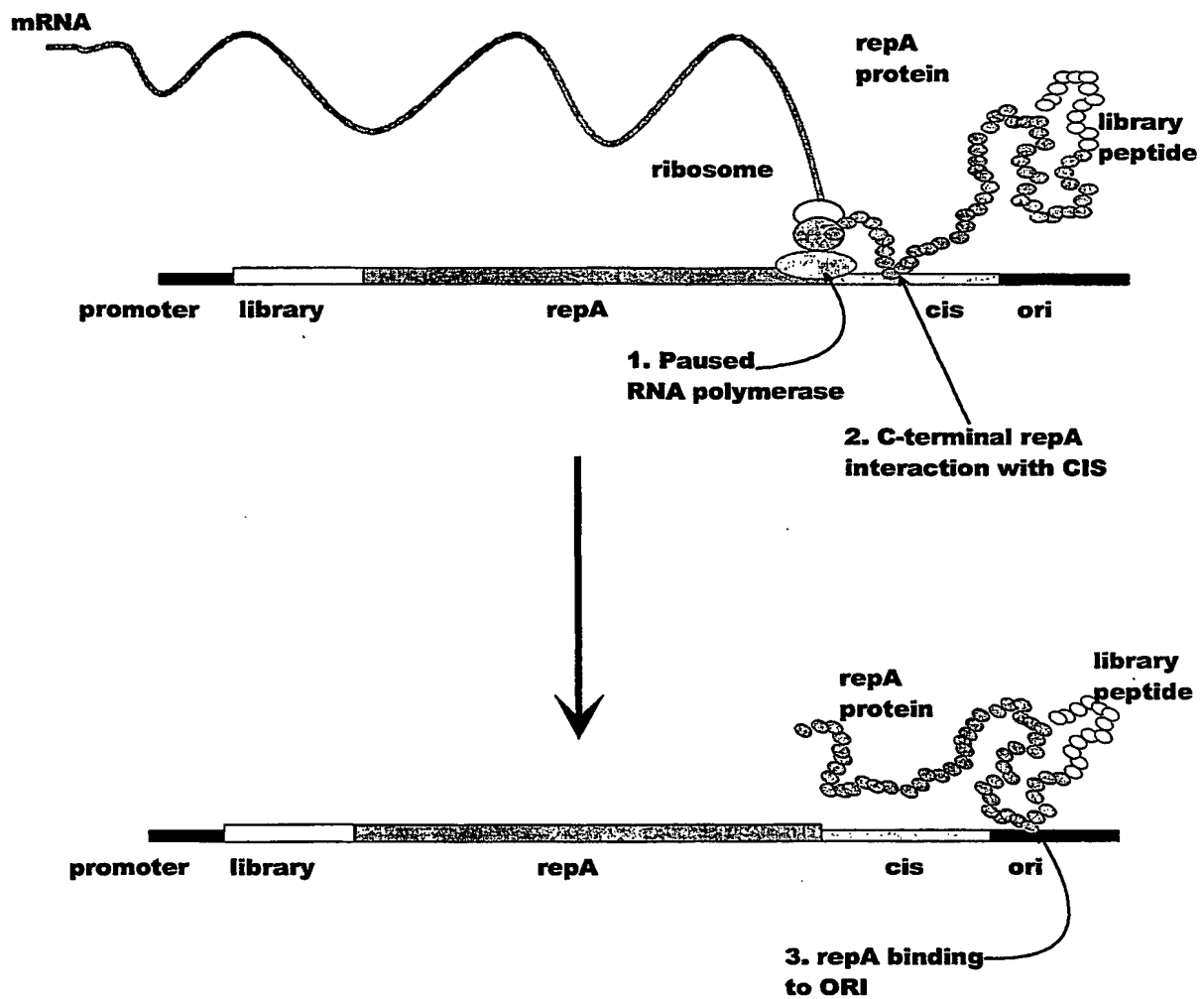


Figure 1

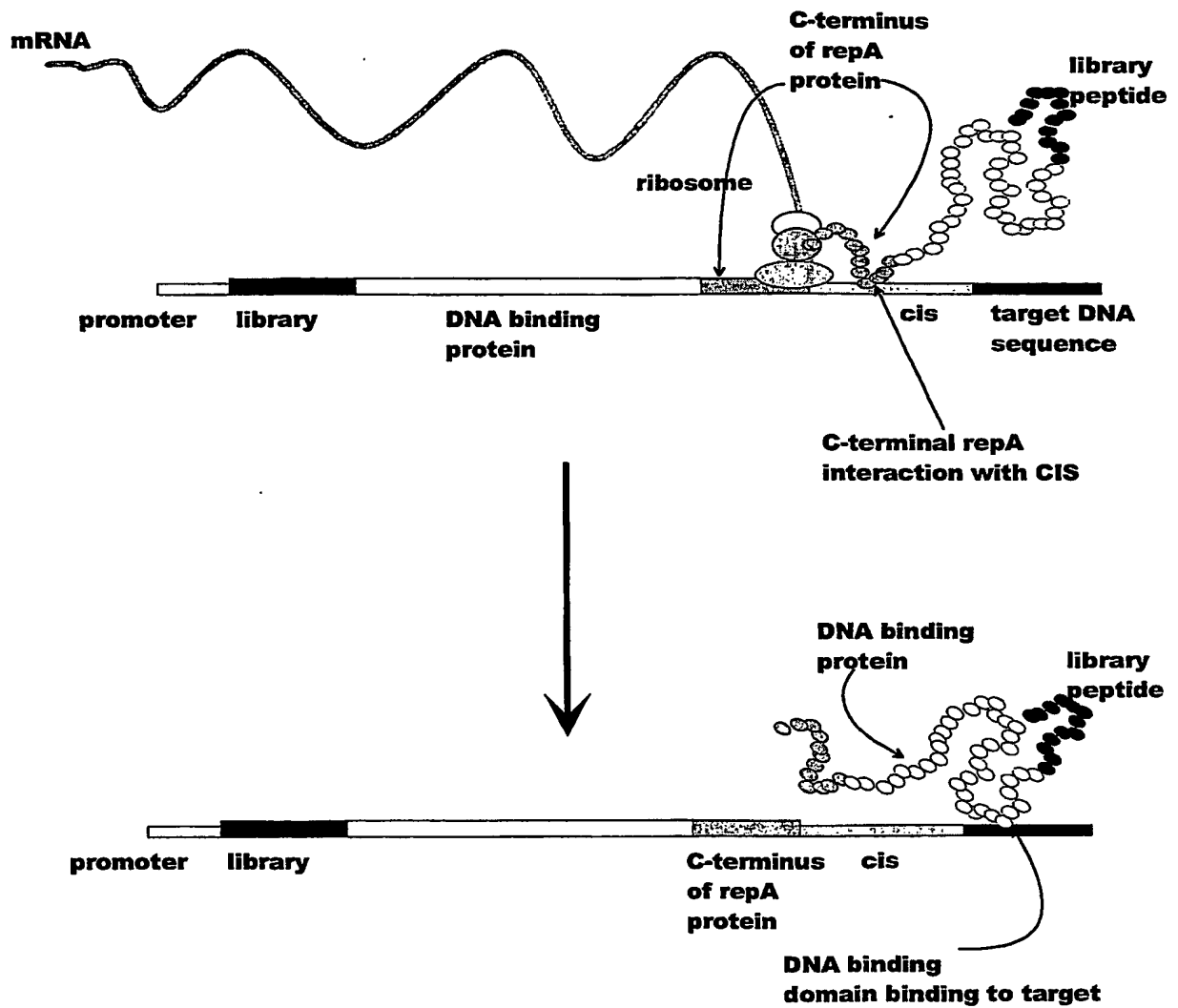
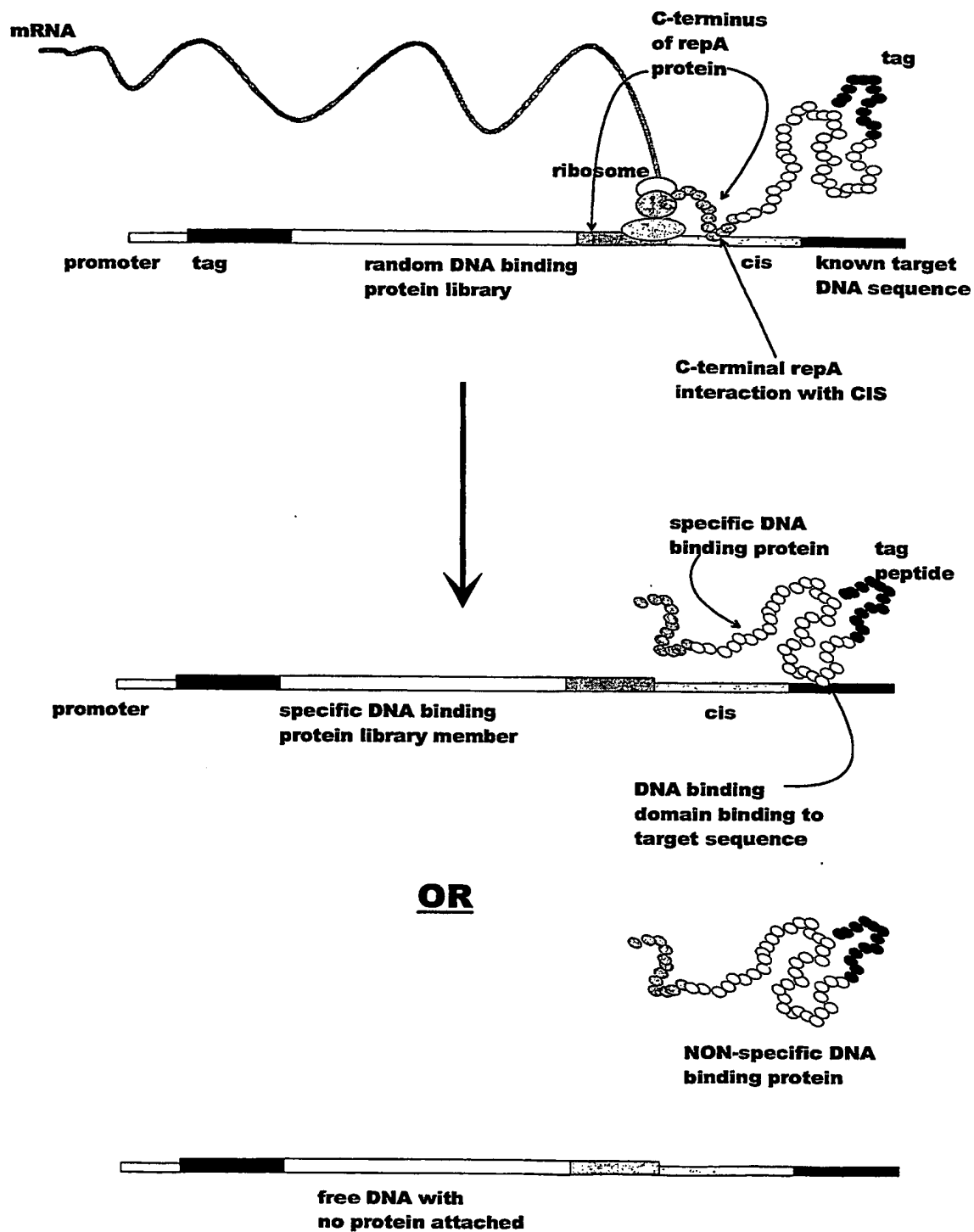
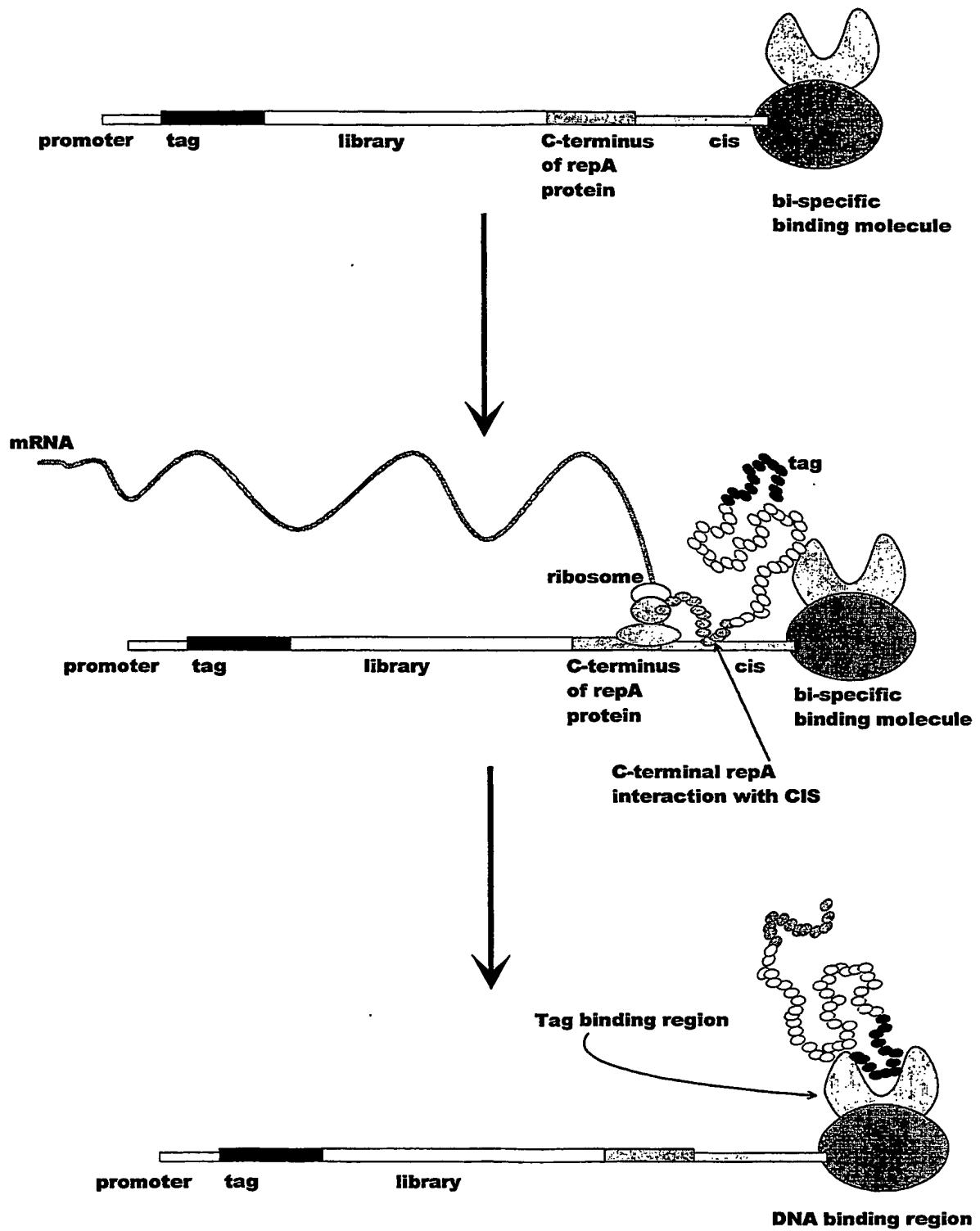


Figure 2



**Figure 3**



**Figure 4**

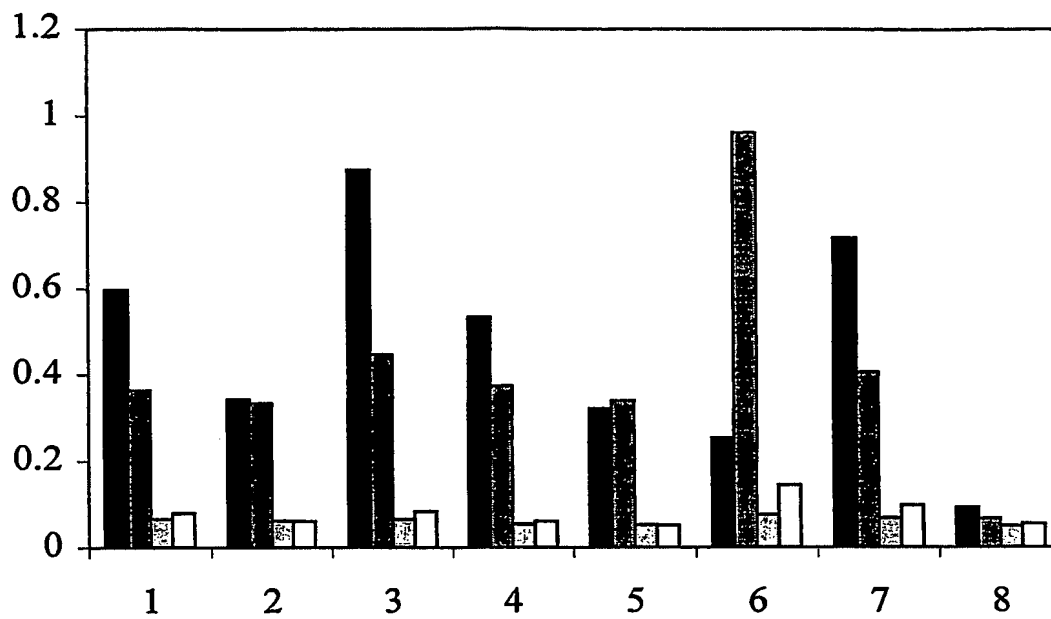


Figure 5

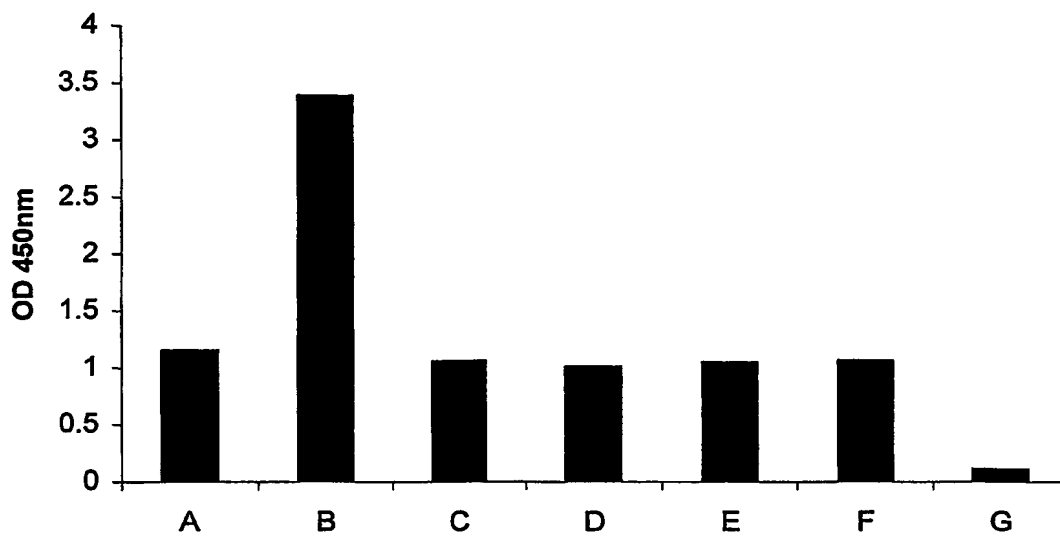


Figure 6

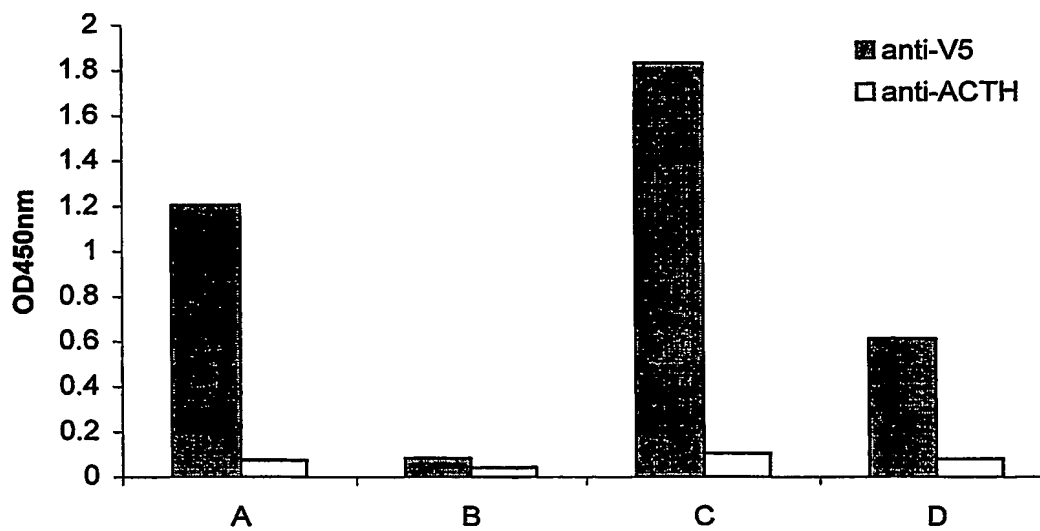


Figure 7